

97	6	5.9	990	1	DPOL_NPVCF	DNA POLYMERASE (EC 2.7	8.75e+00	170	5	5.0	159	1	YXBD_BACSU	HYPOTHETICAL 18.1 KDA	4.58e+02
98	6	5.9	1019	1	ENTK_HUMAN	ENTEROPEPTIDASE PRECUR	8.75e+00	171	5	5.0	159	1	GLBW_CHITP	GLOBIN CTT-W PRECURSOR	4.58e+02
99	6	5.9	1120	1	MPD_RICPR	TRANSCRIPTION-REPAIR C	8.75e+00	172	5	5.0	160	1	YRN5_CAEEL	HYPOTHETICAL 16.4 KDA	4.58e+02
100	6	5.9	1196	1	AMVB_PAEPO	BETA/ALPHA-AMYLASE PRE	8.75e+00	173	5	5.0	161	1	RNKD_PONPY	NONSECRETORY RIBONUCLE	4.58e+02
101	6	5.9	1266	1	SVL_HUMAN	ISOLEUCYL-TRNA SYNTHET	8.75e+00	174	5	5.0	162	1	YNI1_METYL	HYPOTHETICAL PROTEIN I	4.58e+02
102	6	5.9	1313	1	VGLM_PTPV	M POLYPROTEIN PRECURSOR	8.75e+00	175	5	5.0	163	1	PTSN_ECOLI	NITROGEN REGULATORY II	4.58e+02
103	6	5.9	1387	1	RGSC_RAT	REGULATOR OF G-PROTEIN	8.75e+00	176	5	5.0	165	1	Y316_METJA	HYPOTHETICAL PROTEIN R	4.58e+02
104	6	5.9	1475	1	APU_THET	AMYLOPULLULANASE PRECU	8.75e+00	177	5	5.0	171	1	RECX_MYCLE	REGULATORY PROTEIN R	4.58e+02
105	6	5.9	1505	1	SIMA_DROME	SIMILAR PROTEIN.	8.75e+00	178	5	5.0	173	1	Y013_BPPL	PUTATIVE ADENINE-SPECI	4.58e+02
106	6	5.9	3707	1	PGM_MOUSE	BASEMENT MEMBRANE-SPEC	8.75e+00	179	5	5.0	175	1	Y901_MYCTU	HYPOTHETICAL 18.9 KDA	4.58e+02
107	6	5.9	3759	1	TRX_DROME	TRITHORAX PROTEIN.	8.75e+00	180	5	5.0	175	1	DEST_MYXXA	DEVELOPMENT-SPECIFIC P	4.58e+02
108	5	5.0	27	1	K739_PICKL	KILLER TOXIN KT395 (FR	4.58e+02	181	5	5.0	175	1	ATPF_CYACA	ATP SYNTHASE B CHAIN (4.58e+02
109	5	5.0	44	1	RPC3_BPHKO	REGULATORY PROTEIN CII	4.58e+02	182	5	5.0	176	1	COPZ_HUMAN	COATOMER ZETA SUBUNIT	4.58e+02
110	5	5.0	45	1	CSPA_LEIST	MAJOR COLD-SHOCK PROTE	4.58e+02	183	5	5.0	177	1	COPZ_BOVIN	KAPPA CASEIN PRECURSOR	4.58e+02
111	5	5.0	53	1	NCE1_YEAST	NON-CLASSICAL EXPORT P	4.58e+02	184	5	5.0	178	1	CASK_RAT	RIBULOSE BISPHOSPHATE	4.58e+02
112	5	5.0	57	1	CSPB_BACGO	COLD SHOCK PROTEIN CSP	4.58e+02	185	5	5.0	179	1	RBS1_FRIAG	HYPOTHETICAL 19.3 KDA	4.58e+02
113	5	5.0	68	1	CSFC_SALYI	COLD SHOCK-LIKE PROTEI	4.58e+02	186	5	5.0	179	1	Y366_HAEIN	HYPOTHETICAL 19.3 KDA	4.58e+02
114	5	5.0	68	1	CSFC_ECOLI	COLD SHOCK-LIKE PROTEI	4.58e+02	187	5	5.0	183	1	YREP_STRAM	TATA-BOX BINDING PROTE	4.58e+02
115	5	5.0	72	1	YF60_MYCTU	HYPOTHETICAL 8.2 KDA P	4.58e+02	188	5	5.0	183	1	TF2D_METJA	HYPOXANTHINE-GUANINE-X	4.58e+02
116	5	5.0	74	1	IPB2_YEAST	PROTEASE B INHIBITORS	4.58e+02	189	5	5.0	183	1	HGXK_TRIFO	HYPOXANTHINE-GUANINE-X	4.58e+02
117	5	5.0	77	1	YHR2_VACCV	HYPOTHETICAL HOST RANG	4.58e+02	190	5	5.0	185	1	YXAK_BACSU	HYPOXANTHINE-GUANINE-X	4.58e+02
118	5	5.0	85	1	R37A_MYXGL	60S RIBOSOMAL PROTEIN	4.58e+02	191	5	5.0	185	1	YEF_TREPA	ELONGATION FACTOR P (E	4.58e+02
119	5	5.0	87	1	YXGM_ECOLI	HYPOTHETICAL 9.9 KDA P	4.58e+02	192	5	5.0	187	1	YCF4_CYAPA	HYPOTHETICAL 21.2 KDA	4.58e+02
120	5	5.0	89	1	PHMP_ALCEU	PHOSPHOCARRIER PROTEIN	4.58e+02	193	5	5.0	188	1	PAGC_SALTY	COAT PROTEIN (VIRION P	4.58e+02
121	5	5.0	93	1	VSPM_IBVP3	PUTATIVE SMALL MEMBRAN	4.58e+02	194	5	5.0	189	1	COAT_TYMYA	HYPOTHETICAL 21.0 KDA	4.58e+02
122	5	5.0	94	1	ESA6_MYCTU	6 KDA EARLY SECRETORY	4.58e+02	195	5	5.0	190	1	YNR4_YEAST	HYPOTHETICAL 21.0 KDA	4.58e+02
123	5	5.0	96	1	SY01_HUMAN	SMALL INDUCIBLE CYTOKI	4.58e+02	196	5	5.0	191	1	Y222_HAEIN	PUTATIVE FERREDOXIN-LI	4.58e+02
124	5	5.0	107	1	Y453_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	197	5	5.0	193	1	YXGM_HAEIN	PUTATIVE FERREDOXIN-LI	4.58e+02
125	5	5.0	109	1	N016_SOYBN	HYPOTHETICAL 16 PRECURSOR (4.58e+02	198	5	5.0	194	1	PABA_BACSU	PARA-AMINOBENZOATE/ANT	4.58e+02
126	5	5.0	110	1	YHIT_CHLPN	HYPOTHETICAL HIT-LIKE	4.58e+02	199	5	5.0	194	1	PTH_SYNY3	PEPTIDYL-TRNA HYDROLAS	4.58e+02
127	5	5.0	111	1	RNPL_RANCA	RIBONUCLEASE, LIVER (E	4.58e+02	200	5	5.0	195	1	H12_WHEAT	IMIDAZOLEGLYCEROL-PHOS	4.58e+02
128	5	5.0	112	1	YNB7_YEAST	HYPOTHETICAL 12.5 KDA	4.58e+02	201	5	5.0	195	1	VP12_BPHG	MORPHOGENETIC PROTEIN	4.58e+02
129	5	5.0	113	1	HESB_ANASP	HESB PROTEIN.	4.58e+02	202	5	5.0	198	1	YR71_ECOLI	HYPOTHETICAL 21.9 KDA	4.58e+02
130	5	5.0	114	1	YNI0_MYCTU	HYPOTHETICAL 12.8 KDA	4.58e+02	203	5	5.0	200	1	SODM_GANMI	SUPEROXIDE DISMUTASE [4.58e+02
131	5	5.0	121	1	Y121_ECOLI	INSERTION ELEMENT IS2	4.58e+02	204	5	5.0	200	1	ATP4_IPOBA	ATP SYNTHASE DELTA' CH	4.58e+02
132	5	5.0	122	1	Y14K_PMV	HYPOTHETICAL 14.1 KDA	4.58e+02	205	5	5.0	202	1	Y1BF_ECOLI	HYPOTHETICAL 22.6 KDA	4.58e+02
133	5	5.0	123	1	RR13_ODOSI	CHLOROPLAST 30S RIBOSO	4.58e+02	206	5	5.0	203	1	RUVA_ECOLI	HOLLIDAY JUNCTION DNA	4.58e+02
134	5	5.0	124	1	CDX1_RAT	HOMEOBOX PROTEIN CDX-1	4.58e+02	207	5	5.0	203	1	SOAG_CORS1	SARCOSINE OXIDASE GAMM	4.58e+02
135	5	5.0	123	1	YCK3_CHLRE	HYPOTHETICAL 14.4 KDA	4.58e+02	208	5	5.0	203	1	IAAS_HORVU	ALPHA-AMYLASE/SUBTILIS	4.58e+02
136	5	5.0	125	1	YES2_YEAST	HYPOTHETICAL 14.3 KDA	4.58e+02	209	5	5.0	206	1	YXGM_ECOLI	HYPOTHETICAL 22.2 KDA	4.58e+02
137	5	5.0	127	1	GR14_NEOCA	ANTIGEN NC14.1 (FRAGME	4.58e+02	210	5	5.0	211	1	YF78_CAEEL	HYPOTHETICAL 24.3 KDA	4.58e+02
138	5	5.0	128	1	YPRR_ECOLI	HYPOTHETICAL PROTEIN R	4.58e+02	211	5	5.0	213	1	VATO_YEAST	VACUOLAR ATP SYNTHASE	4.58e+02
139	5	5.0	129	1	AZUR_PSEFB	FLAGELLAR BASAL-BODY R	4.58e+02	212	5	5.0	214	1	PITX_RHIME	PUTATIVE PIT ACCESSORY	4.58e+02
140	5	5.0	129	1	FLGB_BORBU	FLAGELLAR BASAL-BODY R	4.58e+02	213	5	5.0	215	1	R56_MYCPN	30S RIBOSOMAL PROTEIN	4.58e+02
141	5	5.0	135	1	YRN6_CAEEL	HYPOTHETICAL 14.2 KDA	4.58e+02	214	5	5.0	216	1	YF22_CAEEL	MINOR FIBRIL SUBUNIT	4.58e+02
142	5	5.0	136	1	YRN6_CAEEL	E6 PROTEIN.	4.58e+02	215	5	5.0	218	1	YF22_CAEEL	HYPOTHETICAL 24.2 KDA	4.58e+02
143	5	5.0	137	1	YF6_BPV1	PHOSPHOLIPASE A2, BASI	4.58e+02	216	5	5.0	218	1	FGFF_MOUSE	FIBROBLAST GROWTH FACT	4.58e+02
144	5	5.0	138	1	PAZB_TRIFL	HYPOTHETICAL 15.0 KDA	4.58e+02	217	5	5.0	219	1	YQBP_BACSU	HYPOTHETICAL 25.3 KDA	4.58e+02
145	5	5.0	139	1	YLF5_CAEEL	HEMOGLOBIN ALPHA-1 CHA	4.58e+02	218	5	5.0	222	1	GPWA_PSEWI	GLUTATHIONE PEROXIDASE	4.58e+02
146	5	5.0	141	1	HBA1_XENBO	HEMOGLOBIN ALPHA-2 CHA	4.58e+02	219	5	5.0	223	1	SODM_ONCVO	SUPEROXIDE DISMUTASE [4.58e+02
147	5	5.0	141	1	HBA2_XENBO	HEMOGLOBIN ALPHA-2 CHA	4.58e+02	220	5	5.0	225	1	YVS7_CAEEL	HYPOTHETICAL 26.0 KDA	4.58e+02
148	5	5.0	143	1	YEN6_YEAST	HYPOTHETICAL 16.6 KDA	4.58e+02	221	5	5.0	225	1	YB67_AERPE	HYPOTHETICAL PROTEIN A	4.58e+02
149	5	5.0	143	1	YIFN_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02	222	5	5.0	226	1	BIOD_MYCTU	DETHIOBIOTIN SYNTHETAS	4.58e+02
150	5	5.0	144	1	GLB3_LAMSP	GIANT HEMOGLOBIN AIII	4.58e+02	223	5	5.0	226	1	BIOD_MYCTO	DETHIOBIOTIN SYNTHETAS	4.58e+02
151	5	5.0	144	1	RL22_MYCGE	50S RIBOSOMAL PROTEIN	4.58e+02	224	5	5.0	227	1	YQVA_METTF	HISTIDINE TRANSPORT SY	4.58e+02
152	5	5.0	144	1	SODM_PARGL	SUPEROXIDE DISMUTASE [4.58e+02	225	5	5.0	228	1	HISQ_ECOLI	HISTIDINE TRANSPORT SY	4.58e+02
153	5	5.0	144	1	PSAH_SPIOL	PHOTOSYSTEM I REACTION	4.58e+02	226	5	5.0	230	1	ALKL_PSEOL	OUTER MEMBRANE PROTEIN	4.58e+02
154	5	5.0	144	1	GLBB_RIPPA	GIANT HEMOGLOBINS B CH	4.58e+02	227	5	5.0	231	1	YHHW_ECOLI	26.3 KDA PROTEIN IN GN	4.58e+02
155	5	5.0	145	1	HS12_CAEEL	HEAT SHOCK PROTEIN HSP	4.58e+02	228	5	5.0	231	1	AQP2_ECOLI	AQUAPORIN 2 (BACTERIAL	4.58e+02
156	5	5.0	146	1	LCRR_YERPS	LOW CALCIUM RESPONSE L	4.58e+02	229	5	5.0	231	1	YUCL_AGRTE	VIRC1 PROTEIN.	4.58e+02
157	5	5.0	146	1	LCRR_YERPS	LOW CALCIUM RESPONSE L	4.58e+02	230	5	5.0	231	1	LUXF_PHOPO	NON-FLUORESCENT FLAVOP	4.58e+02
158	5	5.0	147	1	YV59_CAEEL	HYPOTHETICAL 15.2 KDA	4.58e+02	231	5	5.0	231	1	PRC3_HAEVA	PROBABAL PROTEASOME CO	4.58e+02
159	5	5.0	147	1	CALM_YEAST	CALMODULIN.	4.58e+02	232	5	5.0	236	1	BAC3_CAEEL	BACTERIAL RHODOPSIN CS	4.58e+02
160	5	5.0	148	1	AZUR_PSEAE	AZURIN PRECURSOR.	4.58e+02	233	5	5.0	236	1	ECSC_BACSU	PROTEIN ECSC.	4.58e+02
161	5	5.0	149	1	YG2U_YEAST	HYPOTHETICAL 17.4 KDA	4.58e+02	234	5	5.0	237	1	ATPO_ARATH	ATP SYNTHASE DELTA CHA	4.58e+02
162	5	5.0	151	1	YGS2_ANACE	HYPOTHETICAL PROTEIN I	4.58e+02	235	5	5.0	240	1	Y586_MYCTU	HYPOTHETICAL TRANSCRIP	4.58e+02
163	5	5.0	151	1	CSGA_ECOLI	MAJOR CURLIN SUBUNIT P	4.58e+02	236	5	5.0	241	1	Y293_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02
164	5	5.0	152	1	GLB3_LUCPE	HEMOGLOBIN III (HB III	4.58e+02	237	5	5.0	241	1	PCXB_ACICA	PROTEOCATECHUATE 3,4-DI	4.58e+02
165	5	5.0	153	1	EMX1_HUMAN	HOMEOBOX PROTEIN EMX1	4.58e+02	238	5	5.0	242	1	REIA_CHLTR	RIBOSE 5-PHOSPHATE ISO	4.58e+02
166	5	5.0	155	1	Y721_AGRVI	HYPOTHETICAL 16.3 KDA	4.58e+02	239	5	5.0	243	1	TRPA_CYAME	TRYPTOPHAN SYNTHASE AL	4.58e+02
167	5	5.0	156	1	YMI9_WHEAT	HYPOTHETICAL 18 KDA PR	4.58e+02	240	5	5.0	245	1	YUBG_ECOLI	HYPOTHETICAL 26.3 KDA	4.58e+02
168	5	5.0	156	1	I3MS_BRANA	MICROSPORE-SPECIFIC PR	4.58e+02	241	5	5.0	245	1	PYHD_NPVLD	POLYHEDRIN (MAJOR OCCL	4.58e+02
169	5	5.0	157	1	RS7_BORBU	30S RIBOSOMAL PROTEIN	4.58e+02	242	5	5.0	246	1	RPSB_MYXXA	RNA POLYMERASE SIGMA-B	4.58e+02

243	5	5.0	1	YK01_MYCTU	HYPOTHETICAL 28.7 KDA	4.58e+02	316	1	CCPB_BACSU	CATABOLITE CONTROL PRO	4.58e+02
244	5	5.0	1	AGL4_ARATH	FLORAL HOMEOCTIC PROTEI	4.58e+02	317	1	CDX2_MOUSE	HOMEOBOX PROTEIN CDX-2	4.58e+02
245	5	5.0	1	Y545_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	318	1	MIAA_RICPR	TRNA DELTA(2)-ISOPENTE	4.58e+02
246	5	5.0	1	ABX3_ANTSP	PROBABLE ATP-DEPENDENT	4.58e+02	319	5	JTV1_HUMAN	JTV-1 PROTEIN	4.58e+02
247	5	5.0	1	VMF1_IACKB	MATRIX PROTEIN M1.	4.58e+02	320	5	ATPG_SYNY3	ATP SYNTHASE GAMMA CHA	4.58e+02
248	5	5.0	1	VMF1_IANIL	MATRIX PROTEIN M1.	4.58e+02	321	5	YP76_CAEEL	HYPOTHETICAL 36.7 KDA	4.58e+02
249	5	5.0	1	VMF1_TAFPW	MATRIX PROTEIN M1.	4.58e+02	322	5	VANB_PSEPU	VANILLATE O-DEMETHYLAS	4.58e+02
250	5	5.0	1	VMF1_TALEI	MATRIX PROTEIN M1.	4.58e+02	323	5	ATPG_ANASP	ATP SYNTHASE GAMMA CHA	4.58e+02
251	5	5.0	1	VMF1_TAFOW	MATRIX PROTEIN M1.	4.58e+02	324	5	ATPG_SYNP1	ATP SYNTHASE GAMMA CHA	4.58e+02
252	5	5.0	1	VMF1_TAFOW	MATRIX PROTEIN M1.	4.58e+02	325	5	TRXB_MYCGE	THIOREDOXIN REDUCTASE	4.58e+02
253	5	5.0	1	YC88_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	326	5	TRXB_MYCPN	THIOREDOXIN REDUCTASE	4.58e+02
254	5	5.0	1	KDUD_BACSU	2-DEOXY-D-GLUCONATE 3-	4.58e+02	327	5	ATPG_SYNP6	ATP SYNTHASE GAMMA CHA	4.58e+02
255	5	5.0	1	TRY3_AEDAE	TRYPSIN 3A1 PRECURSOR	4.58e+02	328	5	YZ11_AQUAE	HYPOTHETICAL PROTEIN A	4.58e+02
256	5	5.0	1	GLT1_HAEIN	GLUTAMATE/ASPARTATE TR	4.58e+02	329	5	YIDZ_ECOLI	HYPOTHETICAL TRANSCRIP	4.58e+02
257	5	5.0	1	YABO_HAEIN	HYPOTHETICAL AMINO-ACI	4.58e+02	330	5	PRXC_CALFU	CHLOROPEROXIDASE PRECU	4.58e+02
258	5	5.0	1	Y224_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	331	5	CISY_BARVI	CITRATE SYNTHASE (EC 4	4.58e+02
259	5	5.0	1	YTMN_BACSU	PROBABLE AMINO-ACID AB	4.58e+02	332	5	VP10_RGDV	NONSTRUCTURAL PROTEIN	4.58e+02
260	5	5.0	1	Y96C_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	333	5	CYFS_ECOLI	HYPOTHETICAL 34.6 KDA	4.58e+02
261	5	5.0	1	PANB_SYNY3	PROBABLE 3-METHYL-2-OX	4.58e+02	334	5	CISY_BARBA	CITRATE SYNTHASE (EC 4	4.58e+02
262	5	5.0	1	AQPB_HUMAN	AQUAPORIN 8.	4.58e+02	335	5	IM30_PEA	CHLOROPLAST MEMBRANE-A	4.58e+02
263	5	5.0	1	DHSB_RICPR	SUCCINATE DEHYDROGENAS	4.58e+02	336	5	CAN2_PIG	CALPAIN 2, LARGE ICATA	4.58e+02
264	5	5.0	1	ITVB_CHICK	INTEGRAL MEMBRANE PROT	4.58e+02	337	5	NORZ_BRAJA	MODULATION PROTEIN Z (4.58e+02
265	5	5.0	1	COBB_SALTY	PUTATIVE NICOTINATE-NU	4.58e+02	338	5	CYL_HUMAN	CYTOCHROME C1, HEME PR	4.58e+02
266	5	5.0	1	PRC2_HUMAN	PROTASOME COMPONENT C	4.58e+02	339	5	ASPG_LUPLU	L-ASPARAGINASE (EC 3.5	4.58e+02
267	5	5.0	1	CANS_BOVIN	CALCIUM-DEPENDENT PROT	4.58e+02	340	5	ASPG_LUPAN	L-ASPARAGINASE (EC 3.5	4.58e+02
268	5	5.0	1	YLPK_CAEEL	HYPOTHETICAL 30.3 KDA	4.58e+02	341	5	ASPG_LUPAL	L-ASPARAGINASE (EC 3.5	4.58e+02
269	5	5.0	1	YXEM_BACSU	PROBABLE AMINO-ACID AB	4.58e+02	342	5	YVGR_ECOLI	HYPOTHETICAL 36.0 KDA	4.58e+02
270	5	5.0	1	PRC9_DROME	PROTASOME 29 KDA SUBU	4.58e+02	343	5	MORA_SYNP7	MOLYBDENUM COFACTOR BI	4.58e+02
271	5	5.0	1	OTSB_RHISN	PROBABLE TREHALOSE-PHO	4.58e+02	344	5	COBP_PSESM	COPPER RESISTANCE PROT	4.58e+02
272	5	5.0	1	UBIE_RICPR	PROBABLE UBIQUITONE/ME	4.58e+02	345	5	YHO3_YEAST	HYPOTHETICAL 37.4 KDA	4.58e+02
273	5	5.0	1	ITMB_MOUSE	INTEGRAL MEMBRANE PROT	4.58e+02	346	5	IPNS_STRJU	ISOPENICILLIN N SYNTHE	4.58e+02
274	5	5.0	1	YLI2_MYCHO	HYPOTHETICAL 31.2 KDA	4.58e+02	347	5	YWC8_BACSU	HYPOTHETICAL 36.6 KDA	4.58e+02
275	5	5.0	1	YC73_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02	348	5	DPPE_HAEIN	DIPEPTIDE TRANSPORT SY	4.58e+02
276	5	5.0	1	TRPA_PSESY	TRYPTOPHAN SYNTHASE AL	4.58e+02	349	5	RLAO_SULAC	ACIDIC RIBOSOMAL PROTE	4.58e+02
277	5	5.0	1	Y374_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	350	5	STSL_ARATH	STRICTOSIDINE SYNTHASE	4.58e+02
278	5	5.0	1	YTMK_BACSU	PROBABLE AMINO-ACID AB	4.58e+02	351	5	TMK8_CAEEL	TMK-8 PROTEIN.	4.58e+02
279	5	5.0	1	HIS1_ARATH	IMIDAZOLEGLYCEROL-PHOS	4.58e+02	352	5	HEM1_CHLTR	GLUTAMYL-TRNA REDUCTAS	4.58e+02
280	5	5.0	1	YM35_MYCTU	HYPOTHETICAL 29.8 KDA	4.58e+02	353	5	GLPX_SHIFL	GLPX PROTEIN.	4.58e+02
281	5	5.0	1	YM97_YEAST	HYPOTHETICAL 31.2 KDA	4.58e+02	354	5	GLPX_ECOLI	CINNAMYL-ALCOHOL DEHYD	4.58e+02
282	5	5.0	1	CYL_RHORI	CYTOCHROME C1 PRECURSO	4.58e+02	355	5	CADH_PETCR	HYPOTHETICAL PROTEIN H	4.58e+02
283	5	5.0	1	VG02_BPT4	TERMINAL DNA PROTECTIN	4.58e+02	356	5	YGBQ_HAEIN	KETOL-ACID REDUCTOISOM	4.58e+02
284	5	5.0	1	TRPC_METTM	INDOLE-3-GLYCEROL PHOS	4.58e+02	357	5	ILVC_RHOMO	THREONINE 3-DEHYDROGEN	4.58e+02
285	5	5.0	1	CAPB_MOUSE	F-ACTIN CAPPING PROTEI	4.58e+02	358	5	TDH_XANCP	ALLERGEN MAG (FRAGMENT	4.58e+02
286	5	5.0	1	PRTA_ASPNG	ASPERGILLOPEPSIN II PR	4.58e+02	359	5	MAG_DERFA	BRANCHED-CHAIN AMINO A	4.58e+02
287	5	5.0	1	CXBI_HUMAN	GAP JUNCTION BETA-1 PR	4.58e+02	360	5	Y535_METJA	HYPOTHETICAL PROTEIN M	4.58e+02
288	5	5.0	1	CXBI_RAT	GAP JUNCTION BETA-1 PR	4.58e+02	361	5	ILVE_HAEIN	ARGININE N-SUCCINYLTRA	4.58e+02
289	5	5.0	1	ALKE_BABBO	ALDO-KETO REDUCTASE (F	4.58e+02	362	5	ASTA_ECOLI	NADH-UBIQUINONE OXIDOR	4.58e+02
290	5	5.0	1	ACCD_PORPU	ACETYL-COENZYME A CARB	4.58e+02	363	5	LIMA_PSESS	LIPASE MODULATOR PRECU	4.58e+02
291	5	5.0	1	ILVE_RICPR	PROBABLE BRANCHED-CHAI	4.58e+02	364	5	LIMA_BURCE	HYPOTHETICAL 39.0 KDA	4.58e+02
292	5	5.0	1	MOTV_VIBPA	SODIUM-TYPE FLAGELLAR	4.58e+02	365	5	YJ9S_YEAST	NADH-UBIQUINONE OXIDOR	4.58e+02
293	5	5.0	1	YHAE_ECOLI	HYPOTHETICAL 30.4 KDA	4.58e+02	366	5	NU2M_DIDMA	TETRACYCLINE RESISTANC	4.58e+02
294	5	5.0	1	HK25_CHICK	HOMEOBOX PROTEIN NKX-2	4.58e+02	367	5	TCR_STRRM	INTERFERON REGULATORY	4.58e+02
295	5	5.0	1	YK05_MYCTU	HYPOTHETICAL 30.9 KDA	4.58e+02	368	5	IRF2_HUMAN	THIAZOLE BIOSYNTHETIC	4.58e+02
296	5	5.0	1	SAPR_STRPU	SPERACT PRECURSOR (SPE	4.58e+02	369	5	THI4_ARATH	HYPOTHETICAL 38.5 KDA	4.58e+02
297	5	5.0	1	RN15_YEAST	MRNA 3'-END PROCESSING	4.58e+02	370	5	Y957_METJA	HYPOTHETICAL 38.5 KDA	4.58e+02
298	5	5.0	1	BIEA_HUMAN	BILIVERDIN REDUCTASE A	4.58e+02	371	5	YXAL_BACSU	HYPOTHETICAL 38.5 KDA	4.58e+02
299	5	5.0	1	YHAI_ECOLI	HYPOTHETICAL TRANSCRIP	4.58e+02	372	5	CUP5_GALME	PUPAL CUTICLE PROTEIN	4.58e+02
300	5	5.0	1	PIX3_MOUSE	PITUITARY HOMEOBOX 3 (4.58e+02	373	5	ALF_XANFL	FRUCTOSE-BISPHOSPHATE	4.58e+02
301	5	5.0	1	PIX3_HUMAN	PITUITARY HOMEOBOX 3 (4.58e+02	374	5	ALF2_RHOSH	FRUCTOSE-BISPHOSPHATE	4.58e+02
302	5	5.0	1	MYB3_HORVU	MYB-RELATED PROTEIN HV	4.58e+02	375	5	KLF2_HUMAN	KRUPPEL-LIKE FACTOR 2	4.58e+02
303	5	5.0	1	PTB_CLOAB	PHOSPHATE BUTYRYLTRANS	4.58e+02	376	5	PMFE_PROMI	PUTATIVE MINOR FIBRIA	4.58e+02
304	5	5.0	1	PIX3_RAT	PITUITARY HOMEOBOX 3 (4.58e+02	377	5	DHSO_YEAST	SORBITOL DEHYDROGENASE	4.58e+02
305	5	5.0	1	RDGC_ECOLI	RECOMBINATION ASSOCIAT	4.58e+02	378	5	ATPG_CHLRE	ATP SYNTHASE GAMMA CHA	4.58e+02
306	5	5.0	1	POQB_ACICA	COENZYME PQQ SYNTHESIS	4.58e+02	379	5	TGT_RICPR	QUEUINE TRNA-RIBOSYLTR	4.58e+02
307	5	5.0	1	MOVF_BMV	CELL-TO-CELL MOVEMENT	4.58e+02	380	5	YR62_CAEEL	PUTATIVE SERINE/THREON	4.58e+02
308	5	5.0	1	RNH_BPT4	RIBONUCLEASE H (EC 3.1	4.58e+02	381	5	VDH_STRCO	VALINE DEHYDROGENASE (4.58e+02
309	5	5.0	1	Y4CC_RHISN	PROBABLE DNA-INVERTASE	4.58e+02	382	5	YACI_BACSU	HYPOTHETICAL 41.1 KDA	4.58e+02
310	5	5.0	1	YOPD_YEREN	YOPD PROTEIN.	4.58e+02	383	5	ATPG_SPTOL	ATP SYNTHASE GAMMA CHA	4.58e+02
311	5	5.0	1	YOPD_YERPS	YOPD PROTEIN.	4.58e+02	384	5	NADA_SALTY	QUINOLINATE SYNTHETASE	4.58e+02
312	5	5.0	1	ASPG_LUPAR	L-ASPARAGINASE (EC 3.5	4.58e+02	385	5	YMY7_YEAST	HYPOTHETICAL 42.1 KDA	4.58e+02
313	5	5.0	1	ER25_CANAL	C-4 METHYL STEROL OXID	4.58e+02	386	5	AROF_SCHPO	PUTATIVE PHOSPHO-2-DEH	4.58e+02
314	5	5.0	1	YOH1_AZOVI	HYPOTHETICAL 33.2 KDA	4.58e+02	387	5	Y4WE_RHISN	PROBABLE AMINOTRANSFER	4.58e+02
315	5	5.0	1	ATPG_SPTPL	ATP SYNTHASE GAMMA CHA	4.58e+02	388	5	RF2_STRCO	PEPTIDE CHAIN RELEASE	4.58e+02

389	Y316_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	462	5	5.0	428	DHE4_SYN3	NADP-SPECIFIC GLUTAMAT	4.58e+02
390	ATPG_ODOI	ATP SYNTHASE GAMMA CHA	4.58e+02	463	5	5.0	429	KINB_BACSU	SERULIN KINASE B (4.58e+02
391	YHSC_CLOAB	HYPOTHETICAL 42.4 KDA	4.58e+02	464	5	5.0	430	CISZ_RHLTR	CITRATE SYNTHASE, PLAS	4.58e+02
392	YKQA_CAEEL	HYPOTHETICAL 38.5 KDA	4.58e+02	465	5	5.0	431	YNDP_ECOLI	HYPOTHETICAL 49.6 KDA	4.58e+02
393	3BH1_MESAU	3 BETA-HYDROXYSTEROID	4.58e+02	466	5	5.0	432	AMPD_ECOLI	XAA-PRO AMINOPEPTIDASE	4.58e+02
394	TGT_ECOLI	QUEUINE TRNA-RIBOSYLTR	4.58e+02	467	5	5.0	433	ST11_XENLA	MURCROPEPSIN PRECURSOR	4.58e+02
395	MASP_RAT	MASPIN PRECURSOR (PROT	4.58e+02	468	5	5.0	434	Y181_MYCPN	SERINE/THREONINE-PROTE	4.58e+02
396	WECB_ECOLI	UDP-N-ACETYLGLUCOSAMIN	4.58e+02	469	5	5.0	435	GLYA_HYPM	HYPOTHETICAL PROTEIN M	4.58e+02
397	ACT2_LYTP1	ACTIN, CYTOSKELETAL 2	4.58e+02	470	5	5.0	436	THCB_RHOER	SERINE HYDROXYMETHYLTR	4.58e+02
398	ACTB_STRPU	ACTIN, CYTOSKELETAL II	4.58e+02	471	5	5.0	437	GRDB_CLOLI	CITROCHROME P450 116 (E	4.58e+02
399	ACT1_LYTP1	ACTIN, CYTOSKELETAL IB	4.58e+02	472	5	5.0	438	ENOL_CANAL	GLYCINE REDUCTASE COMP	4.58e+02
400	OE56_NPVAC	OCCCLUSION-DERIVED VIRU	4.58e+02	473	5	5.0	439	COAT_SOCMV	ENOLASE 1 (EC 4.2.1.11	4.58e+02
401	YD5_DROME	TRANSCRIPTION FACTOR D	4.58e+02	474	5	5.0	440	WNT_10A	PROTEIN PRECUR	4.58e+02
402	Y022_NPVOP	HYPOTHETICAL 42.5 KDA	4.58e+02	475	5	5.0	441	ENVZ_HAEIN	HYPOTHETICAL PROTEIN H	4.58e+02
403	Y054_NPVOP	PEPTIDE CHAIN RELEASE	4.58e+02	476	5	5.0	442	TRPC_SALTY	OSMOLARITY SENSOR PROT	4.58e+02
404	RF2_MYCTU	ALPHA-GALACTOSIDASE PR	4.58e+02	477	5	5.0	443	CRQ_DROME	TRYPTOPHAN BIOSYNTHESI	4.58e+02
405	AGAL_COFAR	CYTCHROME B. (HSP40)	4.58e+02	478	5	5.0	444	HEMU_BACSU	HEXOKINASE (EC 2.7.1.1	4.58e+02
406	CYB_CANEA	DNAJ PROTEIN (HSP40)	4.58e+02	479	5	5.0	445	CRQ_DROME	CLUSTERED-ASPARAGINE-R	4.58e+02
407	DNAB_STAAU	GAP JUNCTION ALPHA-1 P	4.58e+02	480	5	5.0	446	TRPC_SALTY	TRYPTOPHAN BIOSYNTHESI	4.58e+02
408	CXAL_BRARE	ENVELOPE POLYPROTEIN G	4.58e+02	481	5	5.0	447	SEL2_CAEEL	INTEGRAL MEMBRANE PROT	4.58e+02
409	ENV_STVM2	ENVELOPE POLYPROTEIN G	4.58e+02	482	5	5.0	448	ROCB_BACSU	ARGININE UTILIZATION R	4.58e+02
410	SSUD_ECOLI	ALKANESULFONATE MONOOX	4.58e+02	483	5	5.0	449	PLSB_PHAVU	GLYCEROL-3-PHOSPHATE A	4.58e+02
411	HPPD_STRAW	4-HYDROXYPHENYLPIRUVAT	4.58e+02	484	5	5.0	450	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
412	YHAD_ECOLI	HYPOTHETICAL 39.1 KDA	4.58e+02	485	5	5.0	451	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
413	Y022_NPVAC	HYPOTHETICAL 43.8 KDA	4.58e+02	486	5	5.0	452	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
414	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	487	5	5.0	453	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
415	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	488	5	5.0	454	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
416	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	489	5	5.0	455	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
417	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	490	5	5.0	456	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
418	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	491	5	5.0	457	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
419	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	492	5	5.0	458	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
420	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	493	5	5.0	459	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
421	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	494	5	5.0	460	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
422	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	495	5	5.0	461	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
423	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	496	5	5.0	462	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
424	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	497	5	5.0	463	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
425	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	498	5	5.0	464	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
426	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	499	5	5.0	465	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
427	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	500	5	5.0	466	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
428	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	501	5	5.0	467	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
429	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	502	5	5.0	468	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
430	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	503	5	5.0	469	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
431	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	504	5	5.0	470	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
432	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	505	5	5.0	471	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
433	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	506	5	5.0	472	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
434	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	507	5	5.0	473	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
435	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	508	5	5.0	474	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
436	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	509	5	5.0	475	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
437	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	510	5	5.0	476	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
438	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	511	5	5.0	477	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
439	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	512	5	5.0	478	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
440	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	513	5	5.0	479	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
441	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	514	5	5.0	480	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
442	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	515	5	5.0	481	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
443	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	516	5	5.0	482	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
444	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	517	5	5.0	483	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
445	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	518	5	5.0	484	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
446	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	519	5	5.0	485	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
447	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	520	5	5.0	486	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
448	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	521	5	5.0	487	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
449	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	522	5	5.0	488	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
450	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	523	5	5.0	489	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
451	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	524	5	5.0	490	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
452	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	525	5	5.0	491	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
453	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	526	5	5.0	492	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
454	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	527	5	5.0	493	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
455	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	528	5	5.0	494	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
456	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	529	5	5.0	495	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
457	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	530	5	5.0	496	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
458	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	531	5	5.0	497	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
459	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	532	5	5.0	498	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
460	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	533	5	5.0	499	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02
461	Y022_NPVAC	HYPOTHETICAL 43.4 KDA	4.58e+02	534	5	5.0	500	YNU6_YEAST	GLUTAMYL-TRNA REDUCTAS	4.58e+02

535	5	5.0	493	1	CETP_HUMAN	CHOLESTERYL ESTER TRAN	4.58e+02	608	5	5.0	562	1	YMX8_YEAST	HYPOTHETICAL 62.6 KDA	4.58e+02
536	5	5.0	493	1	HXK_PLAFA	HEXOKINASE (EC 2.7.1.1)	4.58e+02	609	5	5.0	562	1	HEMA_TAJAP	HAGGLUTININ PRECURSOR	4.58e+02
537	5	5.0	493	1	CETP_MACEA	CHOLESTERYL ESTER TRAN	4.58e+02	610	5	5.0	562	1	EXG2_YEAST	GLUCAN 1,3-BETA-GLUCOS	4.58e+02
538	5	5.0	494	1	Y663_METJA	HYPOTHETICAL PROTEIN M	4.58e+02	611	5	5.0	564	1	VAS5_VAGCC	PROTEIN A55	4.58e+02
539	5	5.0	495	1	UXAA_ECOLI	ALTRONATE HYDROLASE (E	4.58e+02	612	5	5.0	565	1	HEMA_TAGUA	HAGGLUTININ PRECURSO	4.58e+02
540	5	5.0	495	1	HXKG_ASPNC	GLUCOKINASE (EC 2.7.1.	4.58e+02	613	5	5.0	566	1	HEMA_TAKIE	HAGGLUTININ PRECURSO	4.58e+02
541	5	5.0	496	1	KP5_YEAST	PROBABLE RIBOSE-PHOSPH	4.58e+02	614	5	5.0	566	1	HEMA_TALEN	HAGGLUTININ PRECURSO	4.58e+02
542	5	5.0	497	1	SPLI_YEAST	TRNA SPLICING PROTEIN	4.58e+02	615	5	5.0	566	1	HEMA_TAPOE	HAGGLUTININ PRECURSO	4.58e+02
543	5	5.0	497	1	MEK1_YEAST	PROTEIN KINASE MEK1/MR	4.58e+02	616	5	5.0	566	1	NPRE_BACCE	BACILLOLYSIN PRECURSOR	4.58e+02
544	5	5.0	499	1	YP75_METTF	HYPOTHETICAL 54.1 KDA	4.58e+02	617	5	5.0	567	1	PGTA_RAT	RAB GERANYLGERANYLTRAN	4.58e+02
545	5	5.0	499	1	MEP2_YEAST	AMMONIUM TRANSPORTER M	4.58e+02	618	5	5.0	568	1	MET3_ASPTE	SULFATE ADENYLITYLTRANF	4.58e+02
546	5	5.0	500	1	DP51_YEAST	PHOSPHATIDYL SERINE DEC	4.58e+02	619	5	5.0	568	1	RS1_RHIME	30S RIBOSOMAL PROTEIN	4.58e+02
547	5	5.0	500	1	YD91_SCHPO	HYPOTHETICAL PROTEIN C	4.58e+02	620	5	5.0	570	1	HEMA_NDVA	HAGGLUTININ-NEURAMIN	4.58e+02
548	5	5.0	500	1	TACY_CLOPE	PERFERINGOLYSIN O PRECU	4.58e+02	621	5	5.0	571	1	HEMA_NDVH3	HAGGLUTININ-NEURAMIN	4.58e+02
549	5	5.0	500	1	UHPB_SALTY	SENSOR PROTEIN UHPB (E	4.58e+02	622	5	5.0	571	1	PEN3_ADE02	PENTON PROTEIN (VIRION	4.58e+02
550	5	5.0	500	1	CBP3_ORISA	SERINE CARBOXYPEPTIDAS	4.58e+02	623	5	5.0	571	1	PEN3_ADE05	PENTON PROTEIN (VIRION	4.58e+02
551	5	5.0	502	1	PUPP_ECOLI	SODIUM/PROLINE SYMPORT	4.58e+02	624	5	5.0	571	1	HEMA_NDVJ	HAGGLUTININ-NEURAMIN	4.58e+02
552	5	5.0	502	1	DCD2_DROME	AROMATIC-L-AMINO-ACID	4.58e+02	625	5	5.0	571	1	HEMA_NDVI	HAGGLUTININ-NEURAMIN	4.58e+02
553	5	5.0	503	1	YHPE_ECOLI	HYPOTHETICAL ABC TRANS	4.58e+02	626	5	5.0	572	1	ZYX_HUMAN	ZYXIN (ZYXIN 2)	4.58e+02
554	5	5.0	503	1	HSP1_MOUSE	HEAT SHOCK FACTOR PROT	4.58e+02	627	5	5.0	572	1	SYX_HAEIN	PROLYL-TRNA SYNTHETASE	4.58e+02
555	5	5.0	506	1	YCX2_EUGGR	HYPOTHETICAL 59.8 KDA	4.58e+02	628	5	5.0	572	1	MET3_PENCH	SULFATE ADENYLITYLTRANF	4.58e+02
556	5	5.0	506	1	KPY2_YEAST	PYRUVATE KINASE 2 (EC	4.58e+02	629	5	5.0	574	1	VGLF_HRSVR	FUSION GLYCOPROTEIN PR	4.58e+02
557	5	5.0	507	1	IRX3_MOUSE	IROQUOIS-CLASS HOMEODO	4.58e+02	630	5	5.0	574	1	COE2_XENLA	TRANSCRIPTION FACTOR C	4.58e+02
558	5	5.0	508	1	CROC_DROME	FORK HEAD DOMAIN PROTE	4.58e+02	631	5	5.0	575	1	APB3_HUMAN	AMYLLOID BETA A4 PRECUR	4.58e+02
559	5	5.0	509	1	AFGL_YEAST	AFG1 PROTEIN	4.58e+02	632	5	5.0	575	1	COE2_MOUSE	TRANSCRIPTION FACTOR C	4.58e+02
560	5	5.0	510	1	VLI_HPV38	MAJOR CAPSID PROTEIN L	4.58e+02	633	5	5.0	575	1	MLJX_SHEEP	MELATONIN-RELATED RECE	4.58e+02
561	5	5.0	511	1	VGLG_VSVIG	SPIKE GLYCOPROTEIN PRE	4.58e+02	634	5	5.0	576	1	UN87_CAEEL	UNC-87 PROTEIN	4.58e+02
562	5	5.0	511	1	VGLG_VSVJG	SPIKE GLYCOPROTEIN PRE	4.58e+02	635	5	5.0	577	1	HEMA_NDVH4	HAGGLUTININ-NEURAMIN	4.58e+02
563	5	5.0	514	1	YMP8_YEAST	HYPOTHETICAL 59.1 KDA	4.58e+02	636	5	5.0	577	1	HEMA_NDVJ	HAGGLUTININ-NEURAMIN	4.58e+02
564	5	5.0	515	1	INR2_HUMAN	INTERFERON-ALPHA/BETA	4.58e+02	637	5	5.0	579	1	YHVO_YEAST	HYPOTHETICAL 66.1 KDA	4.58e+02
565	5	5.0	515	1	PEPB_MYCTU	PEPTIDASE B (EC 3.-.-	4.58e+02	638	5	5.0	579	1	GPC2_RAT	GLYCIPAN-2 PRECURSOR (4.58e+02
566	5	5.0	516	1	ACHD_BOVIN	ACETYLCHOLINE RECEPTO	4.58e+02	639	5	5.0	579	1	SE10_CAEEL	SEL-10 PROTEIN	4.58e+02
567	5	5.0	518	1	TRCB_XENLA	BETA-TRCP (BETA-TRANS	4.58e+02	640	5	5.0	580	1	UAPC_EMENI	PURINE PERMEASE	4.58e+02
568	5	5.0	518	1	LEU1_BUCRP	2-ISOPROPYLMALATE SYNT	4.58e+02	641	5	5.0	581	1	AMV1_SCHPO	PROBABLE ALPHA-AMYLASE	4.58e+02
569	5	5.0	518	1	ATPA_MYCGA	ATP SYNTHASE ALPHA CHA	4.58e+02	642	5	5.0	584	1	COBA_HUMAN	COMPLEMENT COMPONENT C	4.58e+02
570	5	5.0	521	1	YDA4_MYCTU	HYPOTHETICAL 53.9 KDA	4.58e+02	643	5	5.0	584	1	CNA1_DROME	CAMP-DEPENDENT 3',5'-C	4.58e+02
571	5	5.0	522	1	GLG1_BETVU	GLUCOSE-1-PHOSPHATE AD	4.58e+02	644	5	5.0	586	1	VOID_BPP2	OVERCOMING LYSGENIZAT	4.58e+02
572	5	5.0	525	1	YLD4_CAEEL	HYPOTHETICAL 60.1 KDA	4.58e+02	645	5	5.0	586	1	RRPO_BMYVF	PUTATIVE RNA-DIRECTED	4.58e+02
573	5	5.0	526	1	YHVL_CAEEL	HYPOTHETICAL 59.8 KDA	4.58e+02	646	5	5.0	586	1	ENV_PMPV	ENV POLYPROTEIN PRECUR	4.58e+02
574	5	5.0	527	1	VP5_EHDV1	OUTER CAPSID PROTEIN V	4.58e+02	647	5	5.0	586	1	SYR_STRCO	ARGINYL-TRNA SYNTHETAS	4.58e+02
575	5	5.0	528	1	SERA_MYCTU	D-3-PHOSPHOGLYCERATE D	4.58e+02	648	5	5.0	586	1	HOLI_YEAST	HOLI PROTEIN	4.58e+02
576	5	5.0	528	1	TACY_LISIV	IVANOLYSIN PRECURSOR (4.58e+02	649	5	5.0	586	1	LREI_YEAST	LAMINARASE-RESISTANCE	4.58e+02
577	5	5.0	531	1	YDQ2_SCHPO	HYPOTHETICAL 59.3 KDA	4.58e+02	650	5	5.0	587	1	ENV_SRVI	ENV POLYPROTEIN PRECUR	4.58e+02
578	5	5.0	531	1	TYD2_PAPSO	TYROSINE/DOPA DECARBOX	4.58e+02	651	5	5.0	589	1	RESE_BACSU	SENSOR PROTEIN RESE (E	4.58e+02
579	5	5.0	532	1	ICP0_HSVB	TRANS-ACTING TRANSCRIP	4.58e+02	652	5	5.0	590	1	YNG7_YEAST	HYPOTHETICAL 68.8 KDA	4.58e+02
580	5	5.0	533	1	INV_DEBOC	INVERTASE PRECURSOR (E	4.58e+02	653	5	5.0	590	1	YG28_HSV11	HYPOTHETICAL GENE 28 P	4.58e+02
581	5	5.0	534	1	YM22_CAEEL	HYPOTHETICAL 58.7 KDA	4.58e+02	654	5	5.0	591	1	PPOD_LYCES	POLYPHENOL OXIDASE D P	4.58e+02
582	5	5.0	535	1	YA88_SCHPO	HYPOTHETICAL 60.5 KDA	4.58e+02	655	5	5.0	591	1	SYR_BORBU	ARGINYL-TRNA SYNTHETAS	4.58e+02
583	5	5.0	536	1	YEN1_SCHPO	HYPOTHETICAL 52.9 KDA	4.58e+02	656	5	5.0	593	1	CDAS_BACSH	CYCLOMALTODEXTRINASE (4.58e+02
584	5	5.0	538	1	SCD2_SCHPO	SCD2 PROTEIN	4.58e+02	657	5	5.0	593	1	CN9A_HUMAN	CGMP-SPECIFIC 3',5'-CY	4.58e+02
585	5	5.0	538	1	PYRG_HELPY	CTP SYNTHASE (EC 6.3.4	4.58e+02	658	5	5.0	594	1	RCO3_NEUCR	PROBABLE GLUCOSE TRANS	4.58e+02
586	5	5.0	540	1	HXTD_YEAST	HEXOSE TRANSPORTER HXT	4.58e+02	659	5	5.0	595	1	IF2P_ARCFU	PROBABLE TRANSLATION I	4.58e+02
587	5	5.0	542	1	EAT1_BOVIN	EXCITATORY AMINO ACID	4.58e+02	660	5	5.0	598	1	Y036_HUMAN	HYPOTHETICAL PROTEIN K	4.58e+02
588	5	5.0	543	1	PUR6_PICME	PHOSPHORIBOSYLAMINOIM	4.58e+02	661	5	5.0	599	1	TB22_NEIME	TRANSFERRIN-BINDING PR	4.58e+02
589	5	5.0	543	1	EAT1_RAT	EXCITATORY AMINO ACID	4.58e+02	662	5	5.0	601	1	HMDH_CATRO	3-HYDROXY-3-METHYLGUT	4.58e+02
590	5	5.0	543	1	EAT1_ABTBI	EXCITATORY AMINO ACID	4.58e+02	663	5	5.0	604	1	DEDI_YEAST	PUTATIVE ATP-DEPENDENT	4.58e+02
591	5	5.0	547	1	CH60_KLEPN	ZINC CHAPERONIN (PRO	4.58e+02	664	5	5.0	607	1	PPO_VITVI	POLYPHENOL OXIDASE PRE	4.58e+02
592	5	5.0	548	1	PRZN_RENSA	ZINC METALLOPROTEINASE	4.58e+02	665	5	5.0	608	1	YD6C_SCHPO	HYPOTHETICAL 67.5 KDA	4.58e+02
593	5	5.0	550	1	CBS_HUMAN	CYSTATHIONINE BETA-SYN	4.58e+02	666	5	5.0	609	1	NUSK_HALGR	NADH-UBIQUINONE OXIDOR	4.58e+02
594	5	5.0	550	1	PTR2_HUMAN	PARATHYROID HORMONE RE	4.58e+02	667	5	5.0	613	1	NUSM_POLOR	NADH-UBIQUINONE OXIDOR	4.58e+02
595	5	5.0	551	1	HR38_DROME	PROBABLE NUCLEAR HORMO	4.58e+02	668	5	5.0	614	1	YAZA_SCHPO	HYPOTHETICAL 67.3 KDA	4.58e+02
596	5	5.0	553	1	YMDA_YEAST	HYPOTHETICAL 64.4 KDA	4.58e+02	669	5	5.0	616	1	HEMA_NDVQ	HAGGLUTININ-NEURAMIN	4.58e+02
597	5	5.0	553	1	YIDE_ECOLI	HYPOTHETICAL 58.9 KDA	4.58e+02	670	5	5.0	616	1	HEMA_NDVQ	HAGGLUTININ-NEURAMIN	4.58e+02
598	5	5.0	553	1	MCR4_METVA	METHYL-COENZYME M REDU	4.58e+02	671	5	5.0	616	1	SRE2_CAEEL	SRE-2 PROTEIN	4.58e+02
599	5	5.0	554	1	VGA_BPG4	A AND A* PROTEINS (GPA	4.58e+02	672	5	5.0	616	1	RRPO_PURVI	POTATIVE RNA-DIRECTED	4.58e+02
600	5	5.0	555	1	MIS_MOUSE	MUELLERIAN INHIBITING	4.58e+02	673	5	5.0	620	1	NTTA_HUMAN	SODIUM- AND CHLORIDE-D	4.58e+02
601	5	5.0	556	1	FTHS_CLOCY	FORMATE--TETRAHYDROFOL	4.58e+02	674	5	5.0	621	1	NTTA_MOUSE	SODIUM- AND CHLORIDE-D	4.58e+02
602	5	5.0	557	1	G6P1_KLULA	GLUCOSE-6-PHOSPHATE IS	4.58e+02	675	5	5.0	621	1	NTTA_RAT	SODIUM- AND CHLORIDE-D	4.58e+02
603	5	5.0	557	1	G6P1_MOUSE	GLUCOSE-6-PHOSPHATE IS	4.58e+02	676	5	5.0	622	1	YF54_METJA	HYPOTHETICAL PROTEIN M	4.58e+02
604	5	5.0	558	1	LCB1_YEAST	SERINE PALMITOYLTRANSE	4.58e+02	677	5	5.0	624	1	YGL0_YEAST	HYPOTHETICAL 72.0 KDA	4.58e+02
605	5	5.0	560	1	TF56_SULSH	THERMOPHILIC FACTOR 56	4.58e+02	678	5	5.0	625	1	AMYGL_NEUCR	GLUCOAMYLASE PRECURSOR	4.58e+02
606	5	5.0	560	1	NODU_AZOCA	MODULATION PROTEIN U (4.58e+02	679	5	5.0	625	1	TRF5_YEAST	TOPOISOMERASE 1-RELATE	4.58e+02
607	5	5.0	561	1	MTV1_BACST	MODIFICATION METHYLASE	4.58e+02	680	5	5.0	627	1	YHAB_YEAST	HYPOTHETICAL 70.0 KDA	4.58e+02

681	5	5.0	628	1	RA21_SCHPO	DOUBLE-STRAND-BREAK RE	4.58e+02	754	1	MUTB_MYCTU	PROBABLE METHYLMALONYL	4.58e+02
682	5	5.0	630	1	YD13_SCHPO	HYPOTHETICAL 70.6 KDA	4.58e+02	755	1	RREI_HUMAN	RAS-RESPONSIVE ELEMENT	4.58e+02
683	5	5.0	630	1	Y242_MYCCE	HYPOTHETICAL PROTEIN M	4.58e+02	756	1	YHGF_NEIME	HYPOTHETICAL 83.1 KDA	4.58e+02
684	5	5.0	632	1	VGLG_SYNV	SPIKE GLYCOPROTEIN PRE	4.58e+02	757	1	RRP1_IJWIS	RNA-DIRECTED RNA POLYM	4.58e+02
685	5	5.0	632	1	Y242_MYCPN	HYPOTHETICAL PROTEIN M	4.58e+02	758	1	YKDA_YEAST	HYPOTHETICAL 87.9 KDA	4.58e+02
686	5	5.0	634	1	YCX3_EUGSR	HYPOTHETICAL 78.0 KDA	4.58e+02	759	1	IF39_YEAST	EUKARYOTIC TRANSLATION	4.58e+02
687	5	5.0	635	1	YGC3_YEAST	HYPOTHETICAL 70.6 KDA	4.58e+02	760	1	YHGF_ECOLI	HYPOTHETICAL 87.1 KDA	4.58e+02
688	5	5.0	636	1	YOA4_ECOLI	PROBABLE ATP-DEPENDENT	4.58e+02	761	1	YMA6_YEAST	HYPOTHETICAL 87.1 KDA	4.58e+02
689	5	5.0	636	1	GSH1_HUMAN	GLUTAMATE--CYSTEINE LI	4.58e+02	762	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
690	5	5.0	636	1	PHBC_RHET	POLY-BETA-HYDROXYBUTYR	4.58e+02	763	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
691	5	5.0	638	1	YHES_HAEN	HYPOTHETICAL ABC TRANS	4.58e+02	764	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
692	5	5.0	638	1	GHR_HUMAN	GROWTH HORMONE RECEPTO	4.58e+02	765	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
693	5	5.0	638	1	GHR_RABTI	GROWTH HORMONE RECEPTO	4.58e+02	766	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
694	5	5.0	639	1	NOS2_RHIME	NITROUS-OXIDE REDUCTAS	4.58e+02	767	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
695	5	5.0	639	1	PPO_SPIOI	POLYPHENOL OXIDASE PRE	4.58e+02	768	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
696	5	5.0	641	1	YMPK_PSESY	PATHOGENICITY LOCUS PR	4.58e+02	769	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
697	5	5.0	642	1	YMR3_YEAST	HYPOTHETICAL 73.6 KDA	4.58e+02	770	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
698	5	5.0	643	1	VEL_HPV57	REPLICATION PROTEIN E1	4.58e+02	771	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
699	5	5.0	643	1	VEL_HPV3A	REPLICATION PROTEIN E1	4.58e+02	772	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
700	5	5.0	650	1	GHRH_MOUSE	HIGH MOLECULAR WEIGHT	4.58e+02	773	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
701	5	5.0	650	1	DMR9_MOUSE	DMR-N9 PROTEIN.	4.58e+02	774	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
702	5	5.0	651	1	SEC9_YEAST	PROTEIN TRANSPORT PROT	4.58e+02	775	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
703	5	5.0	651	1	YNE7_YEAST	HYPOTHETICAL 74.8 KDA	4.58e+02	776	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
704	5	5.0	655	1	YK16_YEAST	HYPOTHETICAL 74.7 KDA	4.58e+02	777	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
705	5	5.0	659	1	YK16_YEAST	HYPOTHETICAL 74.7 KDA	4.58e+02	778	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
706	5	5.0	661	1	OAT2_RAT	SODIUM-INDEPENDENT ORG	4.58e+02	779	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
707	5	5.0	662	1	YLN9_CAEEL	HYPOTHETICAL 76.7 KDA	4.58e+02	780	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
708	5	5.0	662	1	ERF2_SCHPO	EUKARYOTIC PEPTIDE CHA	4.58e+02	781	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
709	5	5.0	664	1	PRTP_HSVB2	PROBABLE PROCESSING AN	4.58e+02	782	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
710	5	5.0	665	1	TKT_HAELN	TRANSKETOLASE (EC 2.2.	4.58e+02	783	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
711	5	5.0	669	1	HKG_CONVX	HOST RANGE PROTEIN.	4.58e+02	784	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
712	5	5.0	670	1	NRTC_SYNY3	NITRATE TRANSPORT ATP-	4.58e+02	785	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
713	5	5.0	671	1	CHEA_THEMA	CHEMOTAXIS PROTEIN CHE	4.58e+02	786	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
714	5	5.0	672	1	KPCA_RABIT	PROTEIN KINASE C, ALPH	4.58e+02	787	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
715	5	5.0	672	1	KPCA_RAT	PROTEIN KINASE C, ALPH	4.58e+02	788	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
716	5	5.0	672	1	OSM3_CAEEL	KINESIN-LIKE PROTEIN O	4.58e+02	789	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
717	5	5.0	672	1	KPCA_HUMAN	PROTEIN KINASE C, ALPH	4.58e+02	790	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
718	5	5.0	672	1	KPCA_MOUSE	PROTEIN KINASE C, ALPH	4.58e+02	791	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
719	5	5.0	672	1	KPCA_BOVIN	PROTEIN KINASE C, ALPH	4.58e+02	792	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
720	5	5.0	673	1	Y552_HUMAN	HYPOTHETICAL PROTEIN K	4.58e+02	793	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
721	5	5.0	681	1	YADD_SCHPO	HYPOTHETICAL 74.2 KDA	4.58e+02	794	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
722	5	5.0	683	1	BGH3_HUMAN	TRANSFORMING GROWTH FA	4.58e+02	795	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
723	5	5.0	684	1	LIPE_AERHY	EXTRACELLULAR LIPASE P	4.58e+02	796	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
724	5	5.0	687	1	ILVB_YEAST	ACETOLACTATE SYNTHASE	4.58e+02	797	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
725	5	5.0	687	1	CICK_RABIT	CHLORIDE CHANNEL PROTE	4.58e+02	798	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
726	5	5.0	698	1	YB06_YEAST	HYPOTHETICAL 79.2 KDA	4.58e+02	799	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
727	5	5.0	699	1	K122_STRPU	KINESIN-II 85 KDA SUBU	4.58e+02	800	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
728	5	5.0	700	1	CAN2_CHICK	CALPAIN 2, LARGE [CATA	4.58e+02	801	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
729	5	5.0	701	1	PALL_ORISA	PHENYLALANINE AMMONIA	4.58e+02	802	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
730	5	5.0	702	1	KF3A_HUMAN	KINESIN-LIKE PROTEIN K	4.58e+02	803	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
731	5	5.0	703	1	PCCA_HUMAN	PROPIONYL-COA CARBOXYL	4.58e+02	804	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
732	5	5.0	706	1	YK70_YEAST	HYPOTHETICAL 79.4 KDA	4.58e+02	805	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
733	5	5.0	711	1	RED1_RAT	DOUBLE-STRANDED RNA-SP	4.58e+02	806	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
734	5	5.0	711	1	P022_POJPA	RETROVIRUS-RELATED POL	4.58e+02	807	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
735	5	5.0	712	1	MUTB_RHIME	ENDOGLYCANASE G PRECUR	4.58e+02	808	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
736	5	5.0	712	1	SSK1_YEAST	OSMOLALITY TWO-COMPON	4.58e+02	809	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
737	5	5.0	713	1	YG38_YEAST	HYPOTHETICAL 80.2 KDA	4.58e+02	810	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
738	5	5.0	714	1	CLPB_MYCGE	CLPB PROTEIN.	4.58e+02	811	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
739	5	5.0	714	1	NCPR_CATRO	NADPH-CYTOCHROME P450	4.58e+02	812	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
740	5	5.0	715	1	YHCO_CAEEL	HYPOTHETICAL 81.4 KDA	4.58e+02	813	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
741	5	5.0	721	1	YMW6_YEAST	HYPOTHETICAL 82.0 KDA	4.58e+02	814	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
742	5	5.0	725	1	GUNG_CLOCE	ENDOGLUCANASE G PRECUR	4.58e+02	815	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
743	5	5.0	725	1	AGAL_YEAST	A-AGGLUTININ ATTACHMEN	4.58e+02	816	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
744	5	5.0	725	1	IF39_SCHPO	PROBABLE EUKARYOTIC TR	4.58e+02	817	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
745	5	5.0	728	1	TREB_NEUCR	NEUTRAL TREHALASE (EC	4.58e+02	818	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
746	5	5.0	733	1	YFJ2_YEAST	HYPOTHETICAL 79.7 KDA	4.58e+02	819	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
747	5	5.0	733	1	MUTB_STRCM	METHYLMALONYL-COA MUTA	4.58e+02	820	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
748	5	5.0	742	1	K121_STRPU	KINESIN-II 95 KDA SUBU	4.58e+02	821	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
749	5	5.0	742	1	KM65_YEAST	PROBABLE SERINE/THREON	4.58e+02	822	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
750	5	5.0	742	1	AMSE_ERWAM	AMYLOVRAN BIOSYNTHESI	4.58e+02	823	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
751	5	5.0	743	1	KF3B_MOUSE	KINESIN-LIKE PROTEIN K	4.58e+02	824	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
752	5	5.0	747	1	TREB_EMENI	NEUTRAL TREHALASE (EC	4.58e+02	825	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02
753	5	5.0	748	1	TREB_EMENI	NEUTRAL TREHALASE (EC	4.58e+02	826	1	VP41_ROTSH	OUTER CAPSID PROTEIN V	4.58e+02

827	1	TRAI_STRFR	5	5.0	892	1	TRANSPOSASE FOR TRANSP	4.58e+02	900	5	5.0	1134	1	YML7_YEAST	1	HYPOHETICAL 126.1 KDA	4.58e+02
828	1	YML3_YEAST	5	5.0	899	1	HYPOTHETICAL 103.0 KDA	4.58e+02	901	5	5.0	1147	1	YBFC_HUMAN	1	NARDILYSIN PRECURSOR (4.58e+02
829	1	GUNH_CLOTH	5	5.0	900	1	ENDOGLUCANASE H PRECUR	4.58e+02	902	5	5.0	1152	1	YBFC_HUMAN	1	PROBABILE CATION-TRANSP	4.58e+02
830	1	XPC_MOUSE	5	5.0	901	1	DNA-REPAIR PROTEIN COM	4.58e+02	903	5	5.0	1158	1	R114_HUMAN	1	NUCLEAR FACTOR R1P140	4.58e+02
831	1	MALT_ECOLI	5	5.0	901	1	MALT REGULATORY PROTEI	4.58e+02	904	5	5.0	1159	1	YQPA_CAEEL	1	HYPOTHETICAL 127.4 KDA	4.58e+02
832	1	YB56_METUJ	5	5.0	903	1	CELL DIVISION CYCLE PR	4.58e+02	905	5	5.0	1161	1	NRDC_RAT	1	NARDILYSIN PRECURSOR (4.58e+02
833	1	GLRI_HUMAN	5	5.0	906	1	GLUTAMATE RECEPTOR 1 P	4.58e+02	906	5	5.0	1166	1	ADDB_BACSU	1	ATP-DEPENDENT NUCLEASE	4.58e+02
834	1	NUOG_SALTY	5	5.0	907	1	NADH DEHYDROGENASE I C	4.58e+02	907	5	5.0	1169	1	SUV3_DROME	1	SUPPRESSOR OF VARIEGAT	4.58e+02
835	1	GLRI_RAT	5	5.0	907	1	GLUTAMATE RECEPTOR 1 P	4.58e+02	908	5	5.0	1176	1	SLAP_BACSH	1	SURFACE-LAYER 125 KDA	4.58e+02
836	1	IF42_YEAST	5	5.0	907	1	EUKARYOTIC INITIATION	4.58e+02	909	5	5.0	1180	1	ITAL_RAT	1	INTEGRIN ALPHA-1 PRECU	4.58e+02
837	1	YK62_CAEEL	5	5.0	918	1	HYPOTHETICAL 105.4 KDA	4.58e+02	910	5	5.0	1192	1	METH_MYCTU	1	5-METHYLTETRAHYDROPOLA	4.58e+02
838	1	HXK1_RAT	5	5.0	918	1	HEXOKINASE, TYPE I (EC	4.58e+02	911	5	5.0	1194	1	APAF_HUMAN	1	APOPTOTIC PROTEASE ACT	4.58e+02
839	1	HXK1_MOUSE	5	5.0	918	1	HEXOKINASE, TYPE I (EC	4.58e+02	912	5	5.0	1199	1	RPOC_BACSU	1	DNA-DIRECTED RNA POLYM	4.58e+02
840	1	PMAL_NEUCR	5	5.0	920	1	PLASMA MEMBRANE ATPASE	4.58e+02	913	5	5.0	1202	1	NO33_HUMAN	1	NITRIC-OXIDE SYNTHASE,	4.58e+02
841	1	NIA_CICIN	5	5.0	920	1	NITRATE REDUCTASE (EC	4.58e+02	914	5	5.0	1202	1	SVF_FUGRU	1	VALYL-TRNA SYNTHETASE	4.58e+02
842	1	CN2A_BOVIN	5	5.0	921	1	CGMP-DEPENDENT 3',5'-C	4.58e+02	915	5	5.0	1217	1	YQ99_CAEEL	1	PROBABLE INTEGRIN ALPH	4.58e+02
843	1	NRP_MOUSE	5	5.0	923	1	NEUROFILIN PRECURSOR (4.58e+02	916	5	5.0	1238	1	YQ99_CAEEL	1	HYPOTHETICAL 141.2 KDA	4.58e+02
844	1	NRP_HUMAN	5	5.0	923	1	NEUROFILIN PRECURSOR (4.58e+02	917	5	5.0	1243	1	VG37_BPK3	1	TAIL FIBER PROTEIN GP3	4.58e+02
845	1	CN2A_RAT	5	5.0	928	1	CGMP-DEPENDENT 3',5'-C	4.58e+02	918	5	5.0	1250	1	BXE_CLOBO	1	BOTULINUM NEUROTOXIN T	4.58e+02
846	1	YB91_CAEEL	5	5.0	938	1	HYPOTHETICAL 105.4 KDA	4.58e+02	919	5	5.0	1250	1	BXE_CLOBO	1	BOTULINUM NEUROTOXIN T	4.58e+02
847	1	CN2A_HUMAN	5	5.0	941	1	CGMP-DEPENDENT 3',5'-C	4.58e+02	920	5	5.0	1251	1	PIP5_HUMAN	1	1-PHOSPHATIDYLINOSITOL	4.58e+02
848	1	YLM5_CAEEL	5	5.0	943	1	HYPOTHETICAL 105.9 KDA	4.58e+02	921	5	5.0	1257	1	FLH1_CAEEL	1	FLIGHTLESS-1 PROTEIN H	4.58e+02
849	1	YH19_RHOCA	5	5.0	952	1	HYPOTHETICAL 104.1 KDA	4.58e+02	922	5	5.0	1259	1	YTFN_ECOLI	1	HYPOTHETICAL 136.8 KDA	4.58e+02
850	1	LYAG_MOUSE	5	5.0	953	1	LYSOSOMAL ALPHA-GLUCOS	4.58e+02	923	5	5.0	1291	1	VAC4_HELPY	1	VACUOLATING CYTOTOXIN	4.58e+02
851	1	YC07_YEAST	5	5.0	953	1	HYPOTHETICAL 107.9 KDA	4.58e+02	924	5	5.0	1308	1	ERB4_HUMAN	1	TETANUS TOXIN PRECURS	4.58e+02
852	1	YF59_YEAST	5	5.0	954	1	HYPOTHETICAL 108.7 KDA	4.58e+02	925	5	5.0	1314	1	TETX_CLOTE	1	INTERNAL VIRION PROTEI	4.58e+02
853	1	YF59_YEAST	5	5.0	956	1	HYPOTHETICAL 106.1 KDA	4.58e+02	926	5	5.0	1318	1	VIVD_BPT7	1	TETANUS TOXIN PRECURS	4.58e+02
854	1	YF59_YEAST	5	5.0	958	1	HYPOTHETICAL 108.2 KDA	4.58e+02	927	5	5.0	1331	1	MANB_CALSA	1	BETA-MANNANASE/ENDOGLU	4.58e+02
855	1	VP2_BTIV13	5	5.0	958	1	OUTER CAPSID PROTEIN V	4.58e+02	928	5	5.0	1360	1	GLI1_XENLA	1	ZINC FINGER PROTEIN GL	4.58e+02
856	1	VP2_BTIV3V	5	5.0	959	1	OUTER CAPSID PROTEIN V	4.58e+02	929	5	5.0	1379	1	M3K5_MOUSE	1	MITOGEN-ACTIVATED PROT	4.58e+02
857	1	CHS3_NEUCR	5	5.0	960	1	CHITIN SYNTHASE 3 (EC	4.58e+02	930	5	5.0	1394	1	ITAT_DROME	1	POSITION-SPECIFIC ANTI	4.58e+02
858	1	OMPE_CHLTR	5	5.0	964	1	PUTATIVE OUTER MEMBRAN	4.58e+02	931	5	5.0	1409	1	HAP1_HAEIN	1	ADHESION AND PENETRATI	4.58e+02
859	1	IF3A_YEAST	5	5.0	964	1	EUKARYOTIC TRANSLATION	4.58e+02	932	5	5.0	1427	1	REST_HUMAN	1	RESTIN (CYTOPLASMIC LI	4.58e+02
860	1	SYA_BOMMO	5	5.0	967	1	ALANYL-TRNA SYNTHETASE	4.58e+02	933	5	5.0	1431	1	DAPK_HUMAN	1	DEATH-ASSOCIATED PROTE	4.58e+02
861	1	LONN_CAEEL	5	5.0	971	1	MITOCHONDRIAL LON PROT	4.58e+02	934	5	5.0	1452	1	VGL2_FIPV	1	E2 GLYCOPROTEIN PRECUR	4.58e+02
862	1	COPB_YEAST	5	5.0	973	1	COATOMER BETA SUBUNIT	4.58e+02	935	5	5.0	1458	1	PHLX_RABIT	1	PHOSPHOLIPASE ADPRAB-B	4.58e+02
863	1	AGLU_ASPNG	5	5.0	985	1	ALPHA-GLUCOSIDASE PREC	4.58e+02	936	5	5.0	1466	1	SPAZ_YEAST	1	SPAZ PROTEIN.	4.58e+02
864	1	UVRA_MICLU	5	5.0	992	1	EXCINUCLEASE ABC SUBUN	4.58e+02	937	5	5.0	1481	1	APU_THEET	1	AMYLOPULLULANASE PRECU	4.58e+02
865	1	NISB_MICLU	5	5.0	993	1	NISIN BIOSYNTHESIS PRO	4.58e+02	938	5	5.0	1489	1	YGP0_YEAST	1	HYPOTHETICAL 171.5 KDA	4.58e+02
866	1	DP01_TREPA	5	5.0	997	1	DNA POLYMERASE I (EC 2	4.58e+02	939	5	5.0	1502	1	N170_YEAST	1	NUCLEOPORIN NUP170 (NU	4.58e+02
867	1	YENA_DROME	5	5.0	1002	1	YEMANUCLEIN-ALPHA.	4.58e+02	940	5	5.0	1520	1	GLTB_BACSU	1	GLUTAMATE SYNTHASE [NA	4.58e+02
868	1	SYA_ARATH	5	5.0	1003	1	ALANYL-TRNA SYNTHETASE	4.58e+02	941	5	5.0	1547	1	TOP2_BOMMO	1	DNA TOPOISOMERASE II (4.58e+02
869	1	CINI_YEAST	5	5.0	1014	1	CINI PROTEIN.	4.58e+02	942	5	5.0	1548	1	YJGG_DROME	1	UDP-GLUCOSE:GLYCOPROTE	4.58e+02
870	1	HLXA_ECOLI	5	5.0	1024	1	HEMOLYSIN, PLASMIID.	4.58e+02	943	5	5.0	1549	1	YJW2_YEAST	1	PUTATIVE MEMBRANE GLYC	4.58e+02
871	1	PTP1_CAEEL	5	5.0	1026	1	PROTEIN-TYROSINE PHOSP	4.58e+02	944	5	5.0	1549	1	YIR3_YEAST	1	PUTATIVE MEMBRANE GLYC	4.58e+02
872	1	END1_YEAST	5	5.0	1029	1	VACUOLAR BIOGENESIS PR	4.58e+02	945	5	5.0	1581	1	YGLP_BEV	1	PEPOMER GLYCOPROTEIN	4.58e+02
873	1	MT10_YEAST	5	5.0	1035	1	SULFITE REDUCTASE [NAD	4.58e+02	946	5	5.0	1592	1	YH55_YEAST	1	PROBABLE ATP-DEPENDENT	4.58e+02
874	1	YQ86_CAEEL	5	5.0	1036	1	HYPOTHETICAL 118.2 KDA	4.58e+02	947	5	5.0	1595	1	SOS_DROME	1	SON OF SEVENLESS PROTE	4.58e+02
875	1	YR71_CAEEL	5	5.0	1039	1	HYPOTHETICAL 118.2 KDA	4.58e+02	948	5	5.0	1626	1	TP2B_HUMAN	1	DNA TOPOISOMERASE II,	4.58e+02
876	1	YR71_CAEEL	5	5.0	1039	1	ENDOGLUCANASE/EXOGLUCA	4.58e+02	949	5	5.0	1643	1	RRPO_NMV	1	RNA REPLICATION PROTEI	4.58e+02
877	1	RPOC_WEIHE	5	5.0	1046	1	DNA-DIRECTED RNA POLYM	4.58e+02	950	5	5.0	1651	1	VIT6_CAEEL	1	VITELLOGENIN 6 PRECURS	4.58e+02
878	1	Y032_HUMAN	5	5.0	1050	1	HYPOTHETICAL PROTEIN K	4.58e+02	951	5	5.0	1707	1	194K_TRVSV	1	POTENTIAL 194 KDA PROT	4.58e+02
879	1	RPOC_BACAN	5	5.0	1052	1	DNA-DIRECTED RNA POLYM	4.58e+02	952	5	5.0	1743	1	TAGC_DICDI	1	PRESTALK-SPECIFIC PROT	4.58e+02
880	1	RPOC_LISMU	5	5.0	1053	1	DNA-DIRECTED RNA POLYM	4.58e+02	953	5	5.0	1746	1	TENA_PIG	1	TENASCIN PRECURSOR (TN	4.58e+02
881	1	RPOC_LISIN	5	5.0	1053	1	DNA-DIRECTED RNA POLYM	4.58e+02	954	5	5.0	1763	1	PKSM_BACSU	1	PUTATIVE POLYKETIDE BI	4.58e+02
882	1	RPOC_WEIPA	5	5.0	1054	1	DNA-DIRECTED RNA POLYM	4.58e+02	955	5	5.0	1807	1	VIT2_XENLA	1	VITELLOGENIN A2 PRECUR	4.58e+02
883	1	VP2_AHSV3	5	5.0	1057	1	OUTER CAPSID PROTEIN V	4.58e+02	956	5	5.0	1809	1	TSC2_RAT	1	TUBERIN (TUBEROUS SCLE	4.58e+02
884	1	EG51_XENLA	5	5.0	1060	1	KINESIN-RELATED MOTOR	4.58e+02	957	5	5.0	1868	1	YH00_YEAST	1	HYPOTHETICAL 210.4 KDA	4.58e+02
885	1	YK22_SCHPO	5	5.0	1071	1	HYPOTHETICAL 123.7 KDA	4.58e+02	958	5	5.0	1868	1	YH00_YEAST	1	HYPOTHETICAL 210.4 KDA	4.58e+02
886	1	YK11_YEAST	5	5.0	1077	1	HYPOTHETICAL 119.3 KDA	4.58e+02	959	5	5.0	1868	1	YH00_YEAST	1	HYPOTHETICAL 210.4 KDA	4.58e+02
887	1	HLSS_DROME	5	5.0	1077	1	HAIRLESS PROTEIN.	4.58e+02	960	5	5.0	2016	1	CIN5_HUMAN	1	SODIUM CHANNEL PROTEIN	4.58e+02
888	1	MYL5_CAEEL	5	5.0	1080	1	ISOLEUCINE/THREONINE-PROTE	4.58e+02	961	5	5.0	2029	1	LAR_DROME	1	PROTEIN-TYROSINE PHOSP	4.58e+02
889	1	SYI_TREPA	5	5.0	1091	1	ISOLEUCINE/THREONINE-PROTE	4.58e+02	962	5	5.0	2105	1	POLR_ASGVP	1	GENOME POLYPROTEIN [CO	4.58e+02
890	1	AF17_HUMAN	5	5.0	1093	1	AF-17 PROTEIN.	4.58e+02	963	5	5.0	2110	1	MCAS_MYCBO	1	MYCOCEROSIC ACID SYNTH	4.58e+02
891	1	KOF5_YEAST	5	5.0	1101	1	PROBABLE SERINE/THREON	4.58e+02	964	5	5.0	2136	1	YCF2_MARPO	1	HYPOTHETICAL 259 KDA P	4.58e+02
892	1	YMG6_YEAST	5	5.0	1102	1	HYPOTHETICAL 128.1 KDA	4.58e+02	965	5	5.0	2150	1	SDC3_CAEEL	1	SDC-3 PROTEIN.	4.58e+02
893	1	POL2_RRVVS	5	5.0	1107	1	RNA2 POLYPROTEIN [CONT	4.58e+02	966	5	5.0	2151	1	RRPL_SBOUB	1	RNA POLYMERASE BETA SU	4.58e+02
894	1	RH18_YEAST	5	5.0	1114	1	DNA REPAIR PROTEIN RHC	4.58e+02	967	5	5.0	2204	1	RRPL_NDV8	1	RNA POLYMERASE BETA SU	4.58e+02
895	1	IRE1_YEAST	5	5.0	1115	1	SERINE/THREONINE-PROTE	4.58e+02	968	5	5.0	2209	1	Y166_HUMAN	1	HYPOTHETICAL PROTEIN K	4.58e+02
896	1	CYT4_NEUCR	5	5.0	1117	1	MITOCHONDRIAL PROTEIN	4.58e+02	969	5	5.0	2211	1	PA5_BOVIN	1	COAGULATION FACTOR V P	4.58e+02
897	1	POL_FIVSD	5	5.0	1124	1	POL POLYPROTEIN [CONTA	4.58e+02	970	5	5.0	2225	1	PYRI_MESAU	1	CAD PROTEIN [INCLUDES V	4.58e+02
898	1	POL_FIVPE	5	5.0	1124	1	POL POLYPROTEIN [CONTA	4.58e+02	971	5	5.0	2329	1	YH89_CAEEL	1	HYPOTHETICAL 254.3 KDA	4.58e+02
899	1	BATF3_HUMAN	5	5.0	1132	1	LARGE PROLINE-RICH PRO	4.58e+02	972	5	5.0	2444	1	NTC1_HUMAN	1	NEUROGENIC LOCUS NOTCH	4.58e+02

973 5 5.0 2524 1 NOTC_XENLA NEUROGENIC LOCUS NOTCH 4.58e+02
974 5.0 2531 1 NTCL_MOUSE NEUROGENIC LOCUS NOTCH 4.58e+02
975 5.0 2531 1 NTCL_MOUSE NEUROGENIC LOCUS NOTCH 4.58e+02
976 5.0 2628 1 HAGA_PORGI HEMAGGLUTININ A PRECUR 4.58e+02
977 5.0 2688 1 ZEP1_MOUSE ZINC FINGER PROTEIN 40 4.58e+02
978 5.0 2703 1 NOTC_DROME NEUROGENIC LOCUS NOTCH 4.58e+02
979 5.0 2717 1 ZEP1_HUMAN ZINC FINGER PROTEIN 40 4.58e+02
980 5.0 2748 1 NUM1_YEAST NUCLEAR MIGRATION PROT 4.58e+02
981 5.0 2871 1 FBNI_BOVIN FIBRILLIN 1 PRECURSOR 4.58e+02
982 5.0 3005 1 ZFH2_DROME ZINC-FINGER PROTEIN 2 4.58e+02
983 5.0 3092 1 IRAL_YEAST INHIBITORY REGULATOR P 4.58e+02
984 5.0 3140 1 POLG_PPVRA GENOME POLYPROTEIN [CO 4.58e+02
985 5.0 3140 1 POLG_PPVSK GENOME POLYPROTEIN [CO 4.58e+02
986 5.0 3141 1 POLG_PPVD GENOME POLYPROTEIN [CO 4.58e+02
987 5.0 3144 1 VPI3_YEAST VACUOLAR PROTEIN SORTI 4.58e+02
988 5.0 3206 1 POLG_PSBMV GENOME POLYPROTEIN [CO 4.58e+02
989 5.0 3418 1 BRC2_HUMAN BREAST CANCER TYPE 2 S 4.58e+02
990 5.0 3433 1 UTRP_HUMAN UTRAPHIN (DYSTROPHIN-R 4.58e+02
991 5.0 3587 1 SRFL_BACSU SURFACTIN SYNTHETASE S 4.58e+02
992 5.0 3744 1 YHP9_YEAST HYPOTHETICAL 433.2 KDA 4.58e+02
993 5.0 3924 1 ANK2_HUMAN ANKYRIN 2 (BRAIN ANKYR 4.58e+02
994 5.0 4349 1 DYHC_FUSSO DYNEIN HEAVY CHAIN, CY 4.58e+02
995 5.0 4466 1 DYHC_ANTCR DYNEIN BETA CHAIN, CIL 4.58e+02
996 5.0 4485 1 DYHG_CHLRE DYNEIN GAMMA CHAIN, FL 4.58e+02
997 5.0 4563 1 APB_HUMAN APOLIPOPROTEIN B-100 P 4.58e+02
998 5.0 4568 1 DYHC_CAEEL DYNEIN HEAVY CHAIN, CY 4.58e+02
999 5.0 4660 1 LRP2_RAT LOW-DENSITY LIPOPROTEI 4.58e+02
1000 5.0 4660 1 LRP2_RAT LOW-DENSITY LIPOPROTEI 4.58e+02

ALIGNMENTS

RESULT 1
ID CVN_NOSEL STANDARD; PRT; 101 AA.
AC P81180;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYANOVIRIN-N (CV-N).
OS Nostoc ellipsosporum.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
RN [1]
RP SEQUENCE.
RX MEDLINE; 97445156.
RA Gustafson K.R., Sowder R.C. II, Henderson L.E., Cardellina J.H. II,
RA McMahon J.B., Rajamani U., Pannell L.K., Boyd M.R.;
RA "Isolation, primary sequence determination, and disulfide bond
structure of cyanovirin-N, an anti-HIV (human immunodeficiency virus)
protein from the cyanobacterium Nostoc ellipsosporum.";
RL Biochem. Biophys. Res. Commun. 238:223-228(1997).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE; 99262850.
RA Yang F., Bewley C.A., Louis J.M., Gustafson K.R., Boyd M.R.,
RA Gronenborn A.M., Clore G.M., Wlodawer A.;
RA "Crystal structure of cyanovirin-N, a potent HIV-inactivating protein,
shows unexpected domain swapping.";
RL J. Mol. Biol. 288:403-412(1999).
CC -1- FUNCTION: CAPABLE OF INHIBITING HIV-1 AND HIV-2 INFECTION AND
CC REPLICATION.
CC -1- MISCELLANEOUS: CLEAVAGE OF THE DISULFIDE BONDS RESULTS IN THE LOSS
CC OF ANTI-HIV ACTIVITY.
CC PDB; 3EZM; 23-DEC-98.
KW Antiviral; Protein synthesis inhibitor; 3D-structure.
FT DISULFID 8 22
FT DISULFID 58 73
SQ SEQUENCE 101 AA; 11013 MW; 1F84E5B886CCE973 CRC64;

Query Match 59.4%; Score 60; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 5.03e-177;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 NVDSGLKQPSNFETCRNTQLAGSSELAACKTRAAQGFVSTKINLDDHIANIDGTLKYE 101

QY 42 NVDSGLKQPSNFETCRNTQLAGSSELAACKTRAAQGFVSTKINLDDHIANIDGTLKYE 101
RESULT 2
ID DORS_DROME STANDARD; PRT; 678 AA.
AC P15330;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE EMBRYONIC POLARITY DORSAL PROTEIN.
DL.

GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88042799.
RA Steward R.;
RT "Dorsal, an embryonic polarity gene in Drosophila, is homologous to
the vertebrate proto-oncogene, c-rel.";
RL Science 238:692-694(1987).
RN [2]
RP SUBCELLULAR LOCATION, AND REVISIONS.
RX MEDLINE; 90090617.
RA Steward R.;
RT "Relocalization of the dorsal protein from the cytoplasm to the
nucleus correlates with its function.";
RL Cell 59:1179-1188(1989).

CC -1- FUNCTION: EMBRYONIC DEVELOPMENTAL PROTEIN. THE LATERAL OR VENTRAL
CC PROTEIN IN ITS NUCLEUS DURING THE BLASTODERM STAGE. DORSAL IS A
CC MORPHOGENETIC PROTEIN THAT SPECIFICALLY BINDS TO THE KAPPA B-
CC RELATED CONSENSUS SEQUENCE 5'-GRGAANCC-3', LOCATED IN THE
CC ENHANCER REGION OF ZYGOTIC GENES SUCH AS ZEN, TWIST, SNAIL AND
CC DECAPENTAPLEGIC.
CC -1- SUBCELLULAR LOCATION: IN VENTRAL REGIONS IT IS FIRST CYTOPLASMIC,
CC THEN THE PROTEIN IS RELOCALIZED IN THE NUCLEUS. ITS NUCLEAR
CC LOCALIZATION IS ESSENTIAL TO ITS FUNCTION AS A MORPHOGEN. IN
CC DORSAL REGIONS IT REMAINS CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23702; AAA28479.1; -;
CC PIR; A30350; A30350.
CC HSP; P25799; IBFS.
CC TRANSFAC; T00196; -;
CC FLYBASE; FBgn0000462; d1.
CC PRAM; PF00554; RHD; 1.
CC PRINTS; PR00057; NFKB1NSCPFCT.
CC PROSITE; PS01204; REL_1; 1.
CC Developmental protein; Nuclear protein; Phosphorylation.
KW DOMAIN 47 342 REL-LIKE (RHD).
FT DOMAIN 335 340 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT MOD_RES 312 312 PHOSPHORYLATION (BY CAKP) (POTENTIAL).
SQ SEQUENCE 678 AA; 75475 MW; C23630F44D5FCBAF CRC64;

Query Match 7.9%; Score 8; DB 1; Length 678;
Best Local Similarity 100.0%; Pred. No. 4.79e-04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 SSIDLNSV 188

QY 32 SSIDLNSV 39

```
RESULT 3
ID PRC6_LYCES STANDARD; PRT; 259 AA.
AC O24030;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROTEASOME ALPHA SUBUNIT (EC 3.4.99.46) (MULTICATALYTIC ENDOPEPTIDASE
DE COMPLEX ALPHA SUBUNIT).
GN PSR5.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC Solanaceae; Solanum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LUKULLUS;
RT Ziethe K., Stenzel I., Hertel S.C., Koeck M.;
RT Cloning and characterization of PSR5, a tomato cDNA encoding a 20S
RT subunit from the proteasome repressed by phosphate starvation.*;
RT (In) Plant Gene Register PGR98-065.
CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC PROTEOLYTIC ACTIVITY.
CC -!- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -!- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC ALSO IN THE NUCLEUS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
CC PROTEASOME A-TYPE FAMILY. PROS28 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y14339; CAA74725.1; -
DR HSSP; P25156; LPWA.
DR PFAM; PF00227; proteasome; 1.
DR PROSITE; PS00388; PROTEASOME.A; 1.
DR PROTEASOME; Hydrolase; Protease.
DR SEQUENCE 259 AA; 28480 MW; 700B3638C9F0FB52 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 65 NLDHIA 71
QY 86 NLDHIA 92
RESULT 4
ID MODA_MYCTU STANDARD; PRT; 261 AA.
AC P95157; O05125;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE MOLYBDATE-BINDING PROTEIN PRECURSOR.
GN MODA OR RV1857 OR MTCY359.16C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
```

```
RA Laqueyrie A.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 98295987.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RA complete genome sequence.*;
RL Nature 393:537-544(1998).
CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF MOLYBDENUM INTO THE CELL.
CC PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM MODABCD.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99258; CAA67642.1; -
DR EMBL; Z83859; CAB06130.1; -
DR HSSP; P37329; LAMF.
DR TUBERCULIST; RV1857; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Molybdenum; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 261
FT LIPID 22 22
FT CONFLICT 1 59
FT CONFLICT 100 134
FT CONFLICT 100 134
FT IGCQGGGVAGRSSDKLRHQHGRRCRRQSQEDP (IN
FT REF. 1).
FT SEQUENCE 261 AA; 26576 MW; CFE292F6D595A9F5 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 74 AGSSELA 80
QY 64 AGSSELA 70
RESULT 5
ID BHC1_RHOQ STANDARD; PRT; 291 AA.
AC P47231;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE I (EC 1.13.11.39) (23OHBP
DE OXYGENASE I) (2,3-DIHYDROXYBIPHENYL DIOXYGENASE I) (DHD I).
GN BPHCI.
OS Rhodococcus globerulus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Actinobacteriaceae; Rhodococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P6;
RX MEDLINE; 94171820.
RA Asturias J.A., Eltis L.D., Prucha M., Timmis K.N.;
```

"Analysis of three 2,3-dihydroxybiphenyl 1,2-dioxygenases found in *Rhodococcus globerulus* p6. Identification of a new family of extradiol dioxygenases."; J. Biol. Chem. 269:7807-7815(1994).

CC -|- CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE + H(2)O.

CC -|- COFACTOR: FERROUS ION.

CC -|- PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO BENZOIC ACID AND CHLOROBENZOIC ACIDS.

CC -|- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC	EMBL; X75633; CAA53297.1; --
CD	HSSP; P47228; 1HAN.
DE	PFAM; PF01013; Extradiol_dloxy; 1.
DR	PROSITE; PS00082; EXTRADIOL_DIOXYGENAS; 1.
DT	Oxidoreductase; Dloxygenase; Aromatic hydrocarbons catabolism; Iron.
KW	METAL 146 146
FT	METAL 210 210 IRON (BY SIMILARITY).
FT	METAL 260 260 IRON (BY SIMILARITY).
FT	METAL 291 AA; 32081 MW; 104F189FE1EDDA6A CRC64;
SQ	SEQUENCE

Query Match	6.9%;	Score 7;	DB 1;	Length 291;
Best Local Similarity	100.0%;	pred. No. 8.39e-02;		
Matches	7;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Db	91 SSELAAE 97
QY	66 SSELAAE 72

RESULT 6
ID OSTG_YEAST STANDARD; PRT: 350 AA.
AC P48439;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DOLICHYL-DIPHOSPHOOLIGOSACCHARIDE--PROTEIN GLYCOSYLTRANSFERASE GAMMA
DE SUBUNIT PRECURSOR (EC 2.4.1.119) (OLIGOSACCHARYL TRANSFERASE GAMMA
DE SUBUNIT) (OLIGOSACCHARYL TRANSFERASE 34 KDA SUBUNIT).
GG OST3 OR YOR085W OR YOR3124W.
GC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-37; 99-103; 123-153 & 345-350.
RX MEDLINE; 95348180.
RA Karagözü D., Kelleher D.J., Gilmore R.;
RT "Functional characterization of Ost3p. Loss of the 34-kD subunit of
RT the Saccharomyces cerevisiae oligosaccharyltransferase results in
RT biased underglycosylation of acceptor substrates.";
RL J. Cell Biol. 130:567-577(1995).

[2]
RN
RP
SEQUENCE FROM N.A.
RX
MEDLINE; 97344368.
RA
Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
RA
Schwager C., Paces V., Sander C., Ansoerge W.;
RT
"DNA sequencing and analysis of 130 kb from yeast chromosome XV.";
RL
Yeast 13:655-672(1997).

CC	-!	FUNCTION: MAY ENHANCE OLIGOSACCHARIDE TRANSFER IN VIVO TO A SUBSET OF ACCEPTOR SUBSTRATES.
CC	-!	CATALYTIC ACTIVITY: DOLICHYL DIPHOSPHOOLIGOSACCHARIDE + PROTEIN L-ASPARAGINE = DOLICHYL DIPHOSPHATE + A GLYCOPROTEIN WITH THE OLIGOSACCHARIDE CHAIN ATTACHED BY GLYCOSYLAMINE LINKAGE TO PROTEIN L-ASPARAGINE.
CC		

CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBUNIT: YEAST OST SEEMS TO CONSIST OF SIX DIFFERENT SUBUNITS
CC (ALPHA TO ZETA).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE OST3 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC	EMBL: U25052; AAC49042.1; -
DR	EMBL: X94335; CAA64007.1; -
DR	EMBL: Z74993; CAA99280.1; -
DR	SGD: L0002943; Osm3

Transferase; Endoplasmic reticulum; Transmembrane; Signal.			
KW	1	22	
FT			
SIGNAL	23	350	
CHAIN			
FT			
FT			
FT			
FT			
TRANSMEM	186	203	
TRANSMEM	218	237	
TRANSMEM	272	288	
TRANSMEM	310	329	
SEQUENCE	350 AA:	39483 MW:	FEA573RC1AF90380_CDP64.
SO			

Query Match 6.9%; Score 7; DB 1; Length 350;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Db - 241 RNTQLAG 247
 |||||
 Qv 59 PNTQLAC 65

RESULT	7	STANDARD;	PRT;	440 AA.
AC	PUR8_HELPJ			
AD	Q9ZKA2:			
DT	15-FEB-2000	(Rel. 39, Created)		
DT	15-FEB-2000	(Rel. 39, Last sequence update)		
DT	15-FEB-2000	(Rel. 39, Last annotation update)		
DT	ADENYLOSUCCINATE LYASE (EC 4.3.2.2)	(ADENYLOSUCCINASE) (ASL).		
GN	PURB OR HP1112.			
OS	Helicobacter pylori J99	(Campylobacter pylori J99).		
OC	Bacteria; Proteobacteria;	epsilon subdivision; Helicobacter group;		
OC	Helicobacter.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=J99;			
RC	MEDLINE;	99120557.		
RX	Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,			
RA	Smith D.B., Noonan B., Guild B.C., deJonge B.L., Carmel G.,			
RA	Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,			
RA	Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,			
RA	Trust T.J.;			
RT	"Genomic sequence comparison of two unrelated isolates of the human			
RT	gastric pathogen Helicobacter pylori."			

Nature 397:176-180(1999).
 -1- CATALYTIC ACTIVITY: 1-(5'-PHOSPHORIBOSYL)-4-(N-SUCCINO-CARBOXAMIDE)-5'-AMINOMIDAZOLE + FUMARATE + 5'-PHOSPHORIBOSYL-5'-AMINO-4-IMIDAZOLICARBOXAMIDE (ALSO CATALYZES: N6-(1,2-DICARBOXYETHYL)AMP - FUMARATE + AMP).
 -1- PATHWAY: EIGHT STEP IN DE NOVO PURINE BIOSYNTHESIS.
 -1- SIMILARITY: BELONGS TO THE LYASE 1 FAMILY, ADENYLOSSUCINATE LYASE SUBFAMILY.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF001531; AAD06609.1; -
 DR PFAM: PF00206; lyase_1; 1.
 DR PROSITE: PS00163; FUMARATE_LYASES; 1.
 KW Purine biosynthesis; Lyase.
 SQ SEQUENCE 440 AA; 49867 MW; C67D3C02AEFA2EFE CRC64;

Query Match 6.9%; Score 7; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 327 LNSVIEV 333

QY 36 LNSVIEV 42

IT 8
 ID DNAK_ALCEU STANDARD; PRT; 656 AA.
 AC O33522;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE DNAK PROTEIN (HEAT SHOCK PROTEIN 70) (HSP70).
 GN DNAK.
 OS Alkaligenes eutrophus (Ralstonia eutropha).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH34;
 RA Talbi S., van der Lelie D.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ001727; CAA04955.1; -
 DR HSP; P04475; 1DKX.
 DR PFAM: PF00012; HSP70; 1.
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00297; HSP70.1; 1.
 DR PROSITE: PS00329; HSP70.2; 1.
 DR PROSITE: PS01036; HSP70.3; 1.
 KW Chaperone; ATP-binding; Heat shock.
 SQ SEQUENCE 656 AA; 71300 MW; 5C7D5D2CE22F5F97 CRC64;

Query Match 6.9%; Score 7; DB 1; Length 656;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 377 AIOGSVL 383

QY 12 AIOGSVL 18

RESULT 9
 ID YDBH_ECOLI STANDARD; PRT; 879 AA.
 AC P52645; P77502; P76855;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)

DE HYPOTHETICAL 96.8 KDA PROTEIN IN LDHA-TYNA INTERGENIC REGION.
 OS YDBH.

GN Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;
 RX MEDLINE: 97426617.

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE: 97251357.

RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sanei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).

RN [3]
 RP SEQUENCE OF 1-331 FROM N.A.

RC STRAIN=K12;

RA Bunch P.K., Mat-Jan F., Lee N.A., Deayala B.A., Clark D.P.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [4]
 RP IDENTIFICATION.

RA Rudd K.E.;

RL Unpublished observations (MAR-1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF000235; AAC74463.1; -

DR EMBL: D90776; CAB20902.1; -

DR EMBL: D90777; CAB20910.1; -

DR EMBL: U36928; -; NOT_ANNOTATED_CDS.

DR ECOGENE: EG13180; ydbH.

KW Hypothetical protein.

FT CONFLICT 36 36 I -> L (IN REF. 3).

SQ SEQUENCE 879 AA; 96834 MW; 43892C8391751C1D CRC64;

Query Match 6.9%; Score 7; DB 1; Length 879;

Best Local Similarity 100.0%; Pred. No. 8.39e-02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 801 STKINLD 807

QY 82 STKINLD 88

RESULT 10
 ID PER3_MOUSE STANDARD; PRT; 1113 AA.
 AC O70361;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PERIOD CIRCADIAN PROTEIN 3 (MPER3).
 GN PER3.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=BRAIN;
RX MEDLINE; 98318231.
RA Zylka M.J., Shearman L.P., Weaver D.R., Reppert S.M.;
RT "Three period homologs in mammals: differential light responses in the
RT suprachiasmatic circadian clock and oscillatin transcripts outside of
RT brain."
RL Neuron 20:1103-1110(1998).
CC -!- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
CC TRANSCRIPTIONAL INHIBITION (BY SIMILARITY). EXPRESSION OSCILLATES
CC IN THE SUPRACHIASMATIC NUCLEI (SCN) AND EYES. THE EXPRESSION
CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. EXPRESSED IN HEART, BRAIN,
CC LUNG, LIVER, SKELETAL MUSCLE, TESTIS, AND AT LOW LEVEL IN THE
CC SPLEEN AND KIDNEY. IN BRAIN, MAINLY FOUND IN THE SCN. HIPPOCAMPUS,
CC PIRIFORM CORTEX, AND CEREBELLUM. LOWER LEVEL OF EXPRESSION IN THE
CC NEOCORTEX. EXPRESSION EXHIBITS SYNCHRONOUS OSCILLATIONS IN LIVER,
CC SKELETAL MUSCLE AND TESTIS.
CC -!- INDUCTION: NOT ACUTELY INDUCED BY LIGHT IN SCN DURING SUBJECTIVE
CC NIGHT.
CC -!- SIMILARITY: CONTAINS A PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF050182; AAC40147.1; -
DR MGD; MGI:1277134; PER3.
DR PFAM; PF00989; PAS; 1.
KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
FT DOMAIN 52 86
FT REPEAT 121 187 PAS-1.
FT REPEAT 259 323 PAS-2.
FT DOMAIN 336 376 PAC MOTIF.
FT DOMAIN 562 565 POLY-SER.
SQ SEQUENCE 1113 AA; 120939 MW; 8121E235D100A627 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 1113;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 863 AGSSELA 869
QY 64 AGSSELA 70
RESULT 11
ID ICDB_PSEX STANDARD; PRT; 1148 AA.
AC O30611;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE ICE NUCLEATION PROTEIN.
GN INAK.
OS Pseudomonas syringae.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
CC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KCTC 1832;
RA Jung H.-C., Pan J.-G.;
Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CRYSTALLIZATION IN SUPERCOOLED WATER.
-!- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
-!- DOMAIN: CONTAINS MANY IMPERFECT REPEATS OF THE CONSENSUS
OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHERON A 16-RESIDUE AND A
REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
-!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF013159; AAB66891.1; -
DR PFAM; PF00818; Ice_nucleation; 57.
DR PRINTS; PR00327; ICENUCLEATN
DR PROSITE; PS00314; ICE_NUCLEATION; 38.
KW Ice nucleation; Repeat; Outer membrane.
FT DOMAIN 161 164 POLY-ALA.
FT DOMAIN 180 1099 OCTAPEPTIDE PERIODICITY.
SQ SEQUENCE 1148 AA; 113811 MW; BA4019CF20FAE224 CRC64;
Query Match 6.9%; Score 7; DB 1; Length 1148;
Best Local Similarity 100.0%; Pred. No. 8.39e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 534 QGSVLTS 540
QY 14 QGSVLTS 20
RESULT 12
ID ICBV_PSEX STANDARD; PRT; 1196 AA.
AC O33479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ICE NUCLEATION PROTEIN.
GN INAV.
OS Pseudomonas syringae.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
CC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INA5;
RX MEDLINE; 97462815.
RA Schmid D., Fridmore D., Capitani G., Battistutta R., Neeser J.-R.,
RA Jann A.;
RT "Molecular organisation of the ice nucleation protein Inav from
RT Pseudomonas syringae."
RL FEBS Lett. 414:590-594(1997).
CC -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CRYSTALLIZATION IN SUPERCOOLED WATER.
CC -!- SUBUNIT: MEMBRANE ENVIRONMENT OR AGGREGATION SEEMS TO BE REQUIRED
FOR ICE NUCLEATION ACTIVITY.
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
CC -!- DOMAIN: CONTAINS MANY IMPERFECT REPEATS OF THE CONSENSUS
OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHERON A 16-RESIDUE AND A
REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

FT CONFLICT 229 229 E -> EE (IN REF. 2).
 FT CONFLICT 535 535 V -> M (IN REF. 2).
 FT CONFLICT 546 547 WN -> RT (IN REF. 2).
 FT CONFLICT 574 574 MISSING (IN REF. 2).
 FT CONFLICT 591 591 T -> A (IN REF. 2).
 FT CONFLICT 621 621 T -> N (IN REF. 2).
 FT CONFLICT 1295 1295 V -> L (IN REF. 2).
 SQ SEQUENCE 1416 AA; 158365 MW; 447C8110A775DD42 CRC64;

Query Match 6.9%; Score 7; DB 1; Length 1416;
 Best Local Similarity 100.0%; Pred. No. 8.39e-02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

.Db 527 GSVLTST 533
 |||||
 QY 15 GSVLTST 21

RESULT 15
 ID Y076_NPVAC STANDARD; PRT; 84 AA.
 Q06690;
 01-NOV-1995 (Rel. 32, Created)
 01-NOV-1995 (Rel. 32, Last sequence update)
 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 9.4 KDA PROTEIN IN IAP2-VLF1 INTERGENIC REGION.
 OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE; 94303173.
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Perfer M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 polyhedrosis virus";
 RL Virology 202:586-605(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E2;
 RX MEDLINE; 94172322.
 RA Kool M., Broer R., Zuidema D., Goldbach R.W., Vlak J.M.;
 RT "Nucleotide sequence and genetic organization of a 7.3 kb region (map
 unit 47 to 52.5) of Autographa californica nuclear polyhedrosis virus
 fragment EcoRI-C";
 RL J. Gen. Virol. 75:487-494(1994).
 CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; L22858; AAA66706.1; -;
 EMBL; X71415; CAA50539.1; -;
 PTR; S36691; S36691
 KW Hypothetical protein.
 FT CONFLICT 61 61 S -> N (IN REF. 2).
 FT CONFLICT 71 71 A -> T (IN REF. 2).
 FT CONFLICT 74 74 A -> S (IN REF. 2).
 SQ SEQUENCE 84 AA; 9440 MW; AB9E33C41C914439 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 84;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

.Db 79 NTSSID 84
 |||||
 QY 30 NTSSID 35

RESULT 16
 ID RL44_YEAST STANDARD; PRT; 105 AA.
 AC P02405;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 60S RIBOSOMAL PROTEIN L42 (L44) (YL27) (YP44) (L41).
 GN (RPL42A OR RPL44 OR SCL41A OR RPL41A OR YNL162W OR NL722) AND
 GN (RPL42B OR RPL44 OR SCL41B OR RPL41B OR MAK18 OR YHR141C).
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A. (RPL42A AND RPL42B).
 RX MEDLINE; 92104971.
 RA Kawai S., Murao S., Mochizuki M., Shibuya I., Yano K., Takagi M.;
 RT "Drastic alteration of cycloheximide sensitivity by substitution of
 one amino acid in the L41 ribosomal protein of yeasts";
 RL J. Bacteriol. 174:254-262(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (RPL42A).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE; 96287653.
 RA Nasr F., Becam A.-M., Herbert C.J.;
 RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
 24 complete open reading frames: 18 correspond to new genes, one of
 which encodes a protein similar to the human myotonic dystrophy
 kinase";
 RL Yeast 12:169-175(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (RPL42B).
 RC STRAIN=S288C / AB972;
 RX MEDLINE; 94378003.
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII";
 RL Science 265:2077-2082(1994).
 RN [4]
 RP SEQUENCE.
 RX MEDLINE; 79086263.
 RA Itoh T., Wittmann-Liebold B.;
 RT "The primary structure of protein 44 from the large subunit of yeast
 ribosomes";
 RL FEBS Lett. 96:399-402(1978).
 RN [5]
 RP IDENTIFICATION AS MAK18.
 RX MEDLINE; 95270609.
 RA Carroll K., Wickner R.B.;
 RT "Translation and M1 double-stranded RNA propagation: MAK18 - RPL41B
 and cycloheximide curing";
 RL J. Bacteriol. 177:2887-2891(1995).
 CC -1- PPM: THE RESIDUES AT POSITIONS 40 AND 54 WERE NOT POSITIVELY
 IDENTIFIED BUT ARE CHEMICALLY RELATED TO MONOMETHYLLYSINE.
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L42 IN YEAST.
 CC -1- SIMILARITY: BELONGS TO THE L44E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 EMBL; D10578; BAA01435.1; -;
 EMBL; D10579; BAA01436.1; -;
 DR EMBL; X92517; CAA63277.1; -;

DR EMBL; Z71438; CAA96049.1; ALT_SEQ.

DR EMBL; M62391; AAA34927.1; -.

DR EMBL; M62392; AAA34928.1; -.

DR EMBL; U10398; AAB68420.1; -.

DR PIR; A02781; R6BY44.

DR PIR; B43301; B43301.

DR PIR; C43301; C43301.

DR PIR; S48985; S48985.

DR SGD; L0001731; RPL42A.

DR SGD; L0000990; RPL42B.

DR PFAM; PF00935; Ribosomal_L44; 1.

DR PROSITE; PS01172; RIBOSOMAL_L44E; 1.

KW Ribosomal protein; Cycloheximide resistance; Multigene family.

FT INIT_MET 0

FT VARIANT 55 55 P -> Q (CONFERS RESISTANCE TO

FT CYCLOHEXIMIDE, AN INHIBITOR OF

FT POLYPEPTIDE ELONGATION).

FT CONFLICT 87 88 MISSING (IN REF. 4).

FT SEQUENCE 105 AA; 12080 MW; 916634ECE3CB0121 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 105;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 CKTRAQ 81

QY 73 CKTRAQ 78

RESULT 17

ID HER1_CAEEL STANDARD; PRT; 175 AA.

AC P34704;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-FEB-2000 (Rel. 39, Last annotation update)

DE HER-1 PROTEIN PRECURSOR.

GN HER-1 OR ZK287.8.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE; 93170661.

RA Perry M.D., Li W., Trent C., Robertson B., Fire A., Hageman J.M.,

RA Wood W.B.;

RT "Molecular characterization of the her-1 gene suggests a direct role

in cell signaling during Caenorhabditis elegans sex determination.";

Genes Dev. 7:216-228(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA McMurray A.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: DICTATES MALE DEVELOPMENT. PROBABLY PLAYS A DIRECT ROLE

IN CELL SIGNALING DURING C.ELEGANS SEX DETERMINATION.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- ALTERNATIVE PRODUCTS: TWO TYPES OF TRANSCRIPTS HAVE BEEN

IDENTIFIED: A LARGER; BIOLOGICALLY ACTIVE PRODUCT AND A SMALLER

TRANSCRIPT THAT PRODUCES A NON ACTIVE PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

DR EMBL; Z19595; CAA79650.1; -.

DR EMBL; Z19595; CAA79651.1; -.

DR EMBL; Z70757; CAA94804.1; -.

DR PIR; A46388; A46388.

DR PIR; S32245; S32245.

DR PIR; S32246; S32246.

DR WORMPEP; ZK287.8; CE06617.

KW Developmental protein; Glycoprotein; Signal; Alternative splicing.

FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 175 HER-1 PROTEIN.

FT CARBOHYD 98 98 POTENTIAL.

FT CARBOHYD 163 163 POTENTIAL.

FT VARSPLIC 1 111 MISSING (IN TRUNCATED ISOFORM).

FT SEQUENCE 175 AA; 20172 MW; C330DFE9BB3D869A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 175;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 KINLDD 86

QY 84 KINLDD 89

RESULT 18

ID YH22_VACCV STANDARD; PRT; 184 AA.

AC P17366;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-AUG-1990 (Rel. 15, Last annotation update)

DE HYPOTHETICAL 21.7 KDA HINDIII-C PROTEIN.

OS Vaccinia virus (strain WR)

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89073756.

RA Kotwal G.J., Moss B.;

RT "Analysis of a large cluster of nonessential genes deleted from a

vaccinia virus terminal transposition mutant.";

RL Virology 167:524-537(1988).

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

DR EMBL; M22812; AAA69592.1; -.

DR PIR; A31829; WZVZAL.

KW Hypothetical protein; Early protein.

SQ SEQUENCE 184 AA; 21604 MW; E8CCCCDB5529B293 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 184;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 NTSSID 14

QY 30 NTSSID 35

RESULT 19

ID HTGA_ECOLI STANDARD; PRT; 196 AA.

AC P28697;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE HEAT SHOCK PROTEIN HTGA (HEAT SHOCK PROTEIN HTPY).

GN HTGA OR HTPY.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

RN [1]

RP SEQUENCE FROM N.A.

```

RX MEDLINE; 94003405.
RA James R., Dean D.O., Debbage J.;
RT "Five open reading frames upstream of the dnaK gene of E. coli.";
RL DNA Seq. 3:327-332(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE; 93239687.
RA Missiakas D., Georgopoulos C., Raina S.;
RT "The Escherichia coli heat shock gene htpX: mutational analysis,
cloning, sequencing, and transcriptional regulation.";
RL J. Bacteriol. 175:2613-2624(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 92334977.
RA Yura H., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
Isono K., Microbuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: NOT KNOWN, REQUIRED FOR HIGH-TEMPERATURE GROWTH OF
E. COLI. IT IS POSSIBLE THAT HTGA PROTEIN SOMEHOW REGULATES EITHER
TRANSCRIPTION OF THE RPOH GENE OR THE ACTIVITY OF ITS GENE PRODUCT
SIGMA-32. PLAYS A ROLE OPPOSITE THAT OF DNAJ, DNAJ, AND GREP IN
TERMS OF HEAT SHOCK REGULATION, ANTAGONIZING THE NEGATIVE OF
THESE AND THUS FINE-TUNING THE HEAT SHOCK RESPONSE.
CC -1- INDUCTION: INDUCED AT HIGH TEMPERATURES.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-36 IS THE INITIATOR.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67700; CAA47932.1; ALT INIT.
DR EMBL; L03720; AAA23992.1; -.
DR EMBL; L03720; AAA23993.1; ALT INIT.
DR EMBL; D10483; -, NOT_ANNOTATED_CDS.
DR EMBL; AE000112; AAC73123.1; -.
DR PIR; S28460; S28460.
DR PIR; A40623; A40623.
DR ECGENE; EG11509; HTGA.
KW Heat shock.
FT DOMAIN 57 65 POLY-SER.
SQ SEQUENCE 196 AA; 21225 MW; ECA6154160A40993 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 159 SIDLNS 164
Qy 33 SIDLNS 38
|||||
RESULT 20 STANDARD; PRT; 198 AA.
ID VIF_BIV27
Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 159 SIDLNS 164
Qy 33 SIDLNS 38
|||||
AC PI9563;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VIRION INFECTIVITY FACTOR (Q PROTEIN).
GN VIF.
OS Bovine immunodeficiency virus (isolate 127) (BIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90223985.
RA Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.;
RT "Nucleotide sequence and genome organization of biologically active
proviruses of the bovine immunodeficiency-like virus.";
RL Virology 175:391-409(1990).
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32690; AAA91272.1; -.
DR PIR; C34742; ASLJBT.
DR HIV; M32690; QORF$BIV127.
SQ SEQUENCE 198 AA; 22828 MW; BB9F4E732D52E89F CRC64;
Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 141 SVLTST 146
Qy 16 SVLTST 21
|||||
RESULT 21 STANDARD; PRT; 198 AA.
ID VIF_BIV06
AC PI9562;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE VIRION INFECTIVITY FACTOR (Q PROTEIN).
GN VIF.
OS Bovine immunodeficiency virus (isolate 106) (BIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90223985.
RA Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.;
RT "Nucleotide sequence and genome organization of biologically active
proviruses of the bovine immunodeficiency-like virus.";
RL Virology 175:391-409(1990).
CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32691; -, NOT_ANNOTATED_CDS.
DR HIV; M32691; QORF$BIV106.
SQ SEQUENCE 198 AA; 22827 MW; 119F49C5FB898529 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 159 SIDLNS 164
Qy 33 SIDLNS 38
|||||

```

```
Db 141 SVLTST 146
    |||||
Qy 16 SVLTST 21

RESULT 22
ID GCH1_HAEIN STANDARD; PRT; 218 AA.
AC P43866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE GTP CYCLOHYDROLASE I (EC 3.5.4.16) (GTP-CH-I).
GN FOLE OR H11447.

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RD / KW20;
RA MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae RD."
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: GTP + 2 H(2)O -> FORMATE + 2-AMINO-4-HYDROXY-
CC 6-(ERYTHRO-1,2,3-TRIHYDROXYPROPYL)DIHYDROPTERIDINE TRIPHOSPHATE.
CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF TETRAHYDROFOLATE.
CC -1- SUBUNIT: HOMOPOLYMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U32823; AAC23097.1; -.
LR TIGR; H11447; -.
DR PFAM; PF01227; GTP_cyclohydrol; 1.
DR PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW One-carbon metabolism; Hydrolase; Allosteric enzyme.
FT DISULFID 109 180 BY SIMILARITY.
SQ SEQUENCE 218 AA; 24945 MW; BCC60F0038158D38 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 LTSTCE 110
    |||||
Qy 18 LTSTCE 23

RESULT 23
ID PAA4_ECOLI STANDARD; PRT; 219 AA.
AC P22996;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RESOLVASE.
GN PARA.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=K12;
RA MEDLINE; 92172284.
RA Katzenmeier G., Schmid C., Kellermann J., Lottspeich F., Bacher A.;
RT "Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I
RT from Escherichia coli."
RL Biol. Chem. Hoppe-Seyler 372:991-997(1991).
RN [2]
RN REVISIONS.
RX MEDLINE; 94136208.
RA Schmid C., Meining W., Weinkauff S., Bachmann L., Ritz H.,
RA Eberhardt S., Gimbel W., Werner T., Lahm H.W., Nar H., Bacher A.;
RT "Studies on GTP cyclohydrolase I of Escherichia coli."
RL Adv. Exp. Med. Biol. 338:157-162(1993).
RN [3]
RN SEQUENCE FROM N.A.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=K12;
RX MEDLINE; 92172284.
RA Katzenmeier G., Schmid C., Kellermann J., Lottspeich F., Bacher A.;
RT "Biosynthesis of tetrahydrofolate. Sequence of GTP cyclohydrolase I
RT from Escherichia coli."
RL Biol. Chem. Hoppe-Seyler 372:991-997(1991).
RN [2]
RN REVISIONS.
RX MEDLINE; 94136208.
RA Schmid C., Meining W., Weinkauff S., Bachmann L., Ritz H.,
RA Eberhardt S., Gimbel W., Werner T., Lahm H.W., Nar H., Bacher A.;
RT "Studies on GTP cyclohydrolase I of Escherichia coli."
RL Adv. Exp. Med. Biol. 338:157-162(1993).
RN [3]
RN SEQUENCE FROM N.A.
```


FT SITE 100 INVOLVED IN RECEPTOR BINDING.
 FT SITE 103 INVOLVED IN RECEPTOR BINDING.
 FT DISULFID 89 BY SIMILARITY.
 FT DISULFID 122 BY SIMILARITY.
 FT DISULFID 126 BY SIMILARITY.
 FT DISULFID 116 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 125 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 55 POTENTIAL.
 FT NON_TER 225
 SQ SEQUENCE 225 AA; 25603 MW; 0DAE138B0AA70F0F CRC64;

Query Match 5.9%; Score 6; DB 1; Length 225;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 ABEKTR 92
 |||||
 QY 71 ABEKTR 76

BLT 26
 ID TSIS_SMSAV STANDARD; PRT; 226 AA.
 AC P01128; O41283;
 DT 21-JUL-1986 (Rel. 01. Created)
 DT 21-JUL-1986 (Rel. 01, last sequence update)
 DT 15-JUL-1999 (Rel. 38, last annotation update)
 DE PDGF-RELATED TRANSFORMING PROTEIN P28-SIS.
 GN V-ST5.
 OS Simian sarcoma virus.
 OC Viruses; Retrov. Viruses; Retroviridae; Mammalian type C retroviruses.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 83144004.
 RA Devare S.G., Reddy E.P., Law J.D., Robbins K.C., Aaronson S.A.;
 RT "Nucleotide sequence of the simian sarcoma virus genome:
 RT demonstration that its acquired cellular sequences encode the
 RT transforming gene product p28sis";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:731-735(1983).
 CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

EMBL; V01201; CAA24516.1; ALT_INIT.
 DR PIR; A01381; TVMVSS.
 DR HSP; P01127; 1PDG.
 DR PFAM; PF00341; PDGF; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00249; PDGF_1; 1.
 KW Transforming protein; Oncogene; Growth factor.
 SQ SEQUENCE 226 AA; 25411 MW; A16813ABB95B90C5 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 226;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 ABEKTR 85
 |||||
 QY 71 ABEKTR 76

RESULT 27
 ID FANE_ECOLI STANDARD; PRT; 228 AA.
 AC P25402;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, last sequence update)
 DT 15-JUL-1999 (Rel. 38, last annotation update)
 DE CHAPERONE PROTEIN FANE PRECURSOR.

GN FANE.
 OS Escherichia coli.
 OG Plasmid pFK99.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE B41;
 RX MEDLINE; 91312125.
 RA Bakker D., Vader C.E.M., Roosendaal B., Mooi F.R., Oudega B.,
 RA de Graaf F.K.;
 RT "Structure and function of periplasmic chaperone-like proteins
 RT involved in the biosynthesis of K88 and K99 fimbriae in
 RT enterotoxigenic Escherichia coli.";
 RL Mol. Microbiol. 5:875-886(1991).
 RN [2]
 RP SEQUENCE OF 204-228 FROM N.A.
 RX MEDLINE; 94187244.
 RA Abe N., Moriishi K., Saito M., Naki M.;
 RT "Confirmed nucleotide sequence of fanF of Escherichia coli K99
 RT fimbriae";
 RL Jpn. J. Vet. Res. 41:97-99(1993).
 RN [3]
 RP SEQUENCE OF 207-228 FROM N.A.
 RC STRAIN-ISOLATE B41;
 RX MEDLINE; 91211613.
 RA Simons B.L., Willemssen P.T.J., Bakker D., Roosendaal B.,
 RA de Graaf F.K., Oudega B.;
 RT "Structure, localization and function of FanF, a minor component of
 RT K99 fimbriae of enterotoxigenic Escherichia coli.";
 RL Mol. Microbiol. 4:2041-2050(1990).
 CC -!- FUNCTION: MEDIATES ASSEMBLY OF PILI BY FORMING SOLUBLE MULTIMERIC
 CC COMPLEXES WITH PILI SUBUNITS AS AN INTERMEDIATE STEP IN THE
 CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE PERIPLASMIC PILUS CHAPERONE FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

EMBL; X56001; CAA39474.1; -
 DR EMBL; S70131; AAB30305.1; -
 DR PIR; S12391; S12391.
 DR PFAM; PF00345; pili_assembly; 1.
 DR PRINTS; PR00969; CHAPERONPILI.
 DR PROSITE; PS00635; PILI_CHAPERONE; 1.
 KW Chaperone; Fimbria; Periplasmic; Signal; Immunoglobulin domain;
 KW Plasmid.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 228 CHAPERONE PROTEIN FANE.
 FT DISULFID 157 198 POTENTIAL.
 SQ SEQUENCE 228 AA; 25314 MW; 5F1666C15EEFB949 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 177 SIDLNS 182
 |||||
 QY 33 SIDLNS 38

RESULT 28
 ID NARW_ECOLI STANDARD; PRT; 231 AA.
 AC P19317;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, last sequence update)

```

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RESPIRATORY NITRATE REDUCTASE 2 DELTA CHAIN (EC 1.7.99.4).
GN NARW.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91042410.
RA Blasco F., Iobbi C., Ratouchniak J., Bonnefoy V., Chippaux M.;
RT "Nitrate reductases of Escherichia coli: sequence of the second
FT nitrate reductase and comparison with that encoded by the narGHJI
RL operon.";
RN Mol. Gen. Genet. 222:104-111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RN Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RT Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RN Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-K12;
RX MEDLINE; 97251357.
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RN Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RN Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampel G., Seki Y., Sivasubram S.,
RN Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RN corresponding to the 28.0-40.1 min region on the linkage map.";
RT DNA Res. 3:363-377(1996).
CC -1- FUNCTION: REQUIRED FOR THE ASSEMBLY OF THE NITRATE REDUCTASE-
CC CYTOCHROME B-NR COMPLEX TO BE FULLY ACTIVE IN THE MEMBRANE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: NITRITE + ACCEPTOR = NITRATE + REDUCED
CC ACCEPTOR.
CC -1- SIMILARITY: 69% IDENTITY TO NARJ (THE DELTA CHAIN OF THE FIRST
CC E. COLI NITRATE REDUCTASE ENZYME).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17110; CAA34966.1; -.
CC DR EMBL; AE000243; AAC74548.1; -.
CC DR EMBL; D90786; CAB21022.1; -.
CC DR EMBL; D90787; CAB21036.1; -.
CC DR PIR; S11429; S11429.
CC DR ECGENE; EG10645; NARW
KW Nitrate assimilation; Oxidoreductase.
SQ SEQUENCE 231 AA; 26160 MW; 30AE8EC9AE6290AB CRC64;

Query Match 5.9%; Score 6; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 156 QLAGSS 161
Qy 62 QLAGSS 67

RESULT 29

```

```

ID PDGB_SHEEP STANDARD; PRT; 241 AA.
AC Q95229;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
DE (PDGF-2).
GN PDGFB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-TEXEL; TISSUE-SPLEEN;
RL Woodall C.J., Zhang Z., Watt N.J.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97123; CAA65790.1; -.
CC DR HSSP; P01127; 1PDG.
CC DR PFAM; PF00341; PDGF.
CC DR PRINTS; PR00438; GRCYSKNOT.
CC DR PROSITE; PS00249; PDGF_1; 1.
KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
FT SIGNAL 1 20
FT PROPEP 21 81
FT CHAIN 82 190
FT PROPEP 191 241
FT SITE 108 108
FT SITE 111 111
FT DISULFID 97 141
FT DISULFID 130 178
FT DISULFID 134 180
FT DISULFID 124 124
FT DISULFID 133 133
FT CARBOHYD 63 63
SQ SEQUENCE 241 AA; 27331 MW; 37BE1EC12E7D2863 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
Qy 71 AECKTR 76

RESULT 30
ID PDGB_MOUSE STANDARD; PRT; 241 AA.
AC P31240;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)

```

DE (PDGF-2).
GN PDGFB OR C-SIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 91257844.
RA Bonthron D.T., Sultan P., Collins T.;
RT "Structure of the murine c-sis proto-oncogene (Sis, PDGFB) encoding
RL the B chain of platelet-derived growth factor.";
CC Genomics 10:287-292(1991).
CC
CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M84453; AAA40113.1; JOINED.
CC EMBL; M84448; AAA40113.1; JOINED.
CC EMBL; M84449; AAA40113.1; JOINED.
CC EMBL; M84450; AAA40113.1; JOINED.
CC EMBL; M84451; AAA40113.1; JOINED.
CC EMBL; M84452; AAA40113.1; JOINED.
CC EMBL; M84484; AAA37485.1; JOINED.
CC EMBL; M84845; AAA37485.1; JOINED.
CC EMBL; M84846; AAA37485.1; JOINED.
CC EMBL; M84847; AAA37485.1; JOINED.
CC EMBL; M84848; AAA37485.1; JOINED.
CC PIR; A39073; PFMSG8.
CC HSP; P01127; 1PDG.
CC MGD; MG1-97528; PDGFB.
CC PFAM; PF00341; PDGF; 1.
CC PRINTS; PF00438; GFCYSKNOT.
CC PROSITE; PS00249; PDGF_1; 1.
CC Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
CC SIGNAL 1 20
CC PROPEP 21 81
CC CHAIN 82 190
CC PROPEP 191 241
CC SITE 108 108
CC SITE 111 111
CC SITE 111 111
CC DISULFID 97 141
CC DISULFID 130 178
CC DISULFID 134 180
CC DISULFID 124 124
CC DISULFID 133 133
CC CARBOHYD 63 63
CC SEQUENCE 241 AA; 27381 MW; 3C5EB7A2DAD64178 CRC64;
CC
CC Query Match 5.9%; Score 6; DB 1; Length 241;
CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Db 95 AECKTR 100
CC |||||
CC Qy 71 AECKTR 76

RESULT 31
ID PDGFB_HUMAN STANDARD; PRT; 241 AA.
AC P01127; P78431;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-FEB-2000 (Rel. 39, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
DE (PDGF-2) (BECAPLERMIN).
GN PDGFB OR C-SIS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 84250225.
RA Josephs S.F., Ratner L., Clarke M.F., Westin E.H., Reitz M.S.,
RA Wong-Staal F.;
RT "Transforming potential of human c-sis nucleotide sequences encoding
RT platelet-derived growth factor.";
RL Science 225:636-639(1984).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 86205961.
RA Rao C.D., Igarashi H., Chiu I.-M., Robbins K.C., Aaronson S.A.;
RT "Structure and sequence of the human c-sis/platelet-derived growth
RT factor 2 (SIS/PDGF2) transcriptional unit.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:2392-2396(1986).
RN [3]
RN SEQUENCE OF 22-241 FROM N.A.
RX MEDLINE; 84205633.
RA Chiu I.-M., Reddy E.P., Givol D., Robbins K.C., Tronick S.R.,
RA Aaronson S.A.;
RT "Nucleotide sequence analysis identifies the human c-sis
RT proto-oncogene as a structural gene for platelet-derived growth
RT factor.";
RL Cell 37:123-129(1984).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE; 85296313.
RA Collins T., Ginsburg D., Boss J.M., Orkin S.H., Pober J.S.;
RT "Cultured human endothelial cells express platelet-derived growth
RT factor B chain: cDNA cloning and structural analysis.";
RL Nature 316:748-750(1985).
RN [5]
RN SEQUENCE FROM N.A.
RX MEDLINE; 85269623.
RA Ratner L., Josephs S.F., Jarrett R., Reitz M.S., Wong-Staal F.;
RT "Nucleotide sequence of transforming human c-sis cDNA clones with
RT homology to platelet-derived growth factor.";
RL Nucleic Acids Res. 13:5007-5018(1985).
RN [6]
RN SEQUENCE FROM N.A.
RX MEDLINE; 87217119.
RA Rao C.D., Igarashi H., Pech M.W., Robbins K.C., Aaronson S.A.;
RT "Oncogenic potential of the human platelet-derived growth factor
RT transcriptional unit.";
RL Cold Spring Harb. Symp. Quant. Biol. 51:959-966(1986).
RN [7]
RN SEQUENCE FROM N.A.
RX MEDLINE; 97141927.
RA Simon M.-P., Pedetour F., Sirvent N., Grosgeorge J., Minoletti F.,
RA Coindre J.-M., Terrier-Lacombe M.-J., Mandahl N., Craver R.D.,
RA Blin N., Sozzi G., Turc-Carel C., O'Brien K.P., Kedra D.,
RA Franssion I., Guilbaud C., Dumanski J.P.;
RT "Deregulation of the platelet-derived growth factor B-chain gene via
RT fusion with collagen gene COL1A1 in dermatofibrosarcoma protuberans
RT and giant-cell fibroblastoma.";
RL Nat. Genet. 15:95-98(1997).

[9] SEQUENCE OF 26-241 FROM N.A.
 RX MEDLINE; 86184981.
 RA "The human osteosarcoma cell line U-2 OS expresses a 3.8 kilobase
 RT mRNA which codes for the sequence of the PDGF-B chain.";
 RL FEBS Lett. 198:344-348(1986).
 RN [10]
 RP SEQUENCE OF 153-200 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE; 84236121.
 RA Johansson A., Heldin C.H., Westesson A., Westermark B., Deuel T.F.,
 RA Huang J.S., Seeburg P.H., Gray A., Ullrich A., Strace G.,
 RA Stroobant P., Waterfield M.D.;
 RT "The c-sis gene encodes a precursor of the B chain of
 RL platelet-derived growth factor.";
 RN EMBO J. 3:921-928(1984).
 RN [11]
 RP SEQUENCE OF 82-110.
 RX MEDLINE; 83197379.
 RA Antoniadou H.N., Hunkapiller M.W.;
 RA "Human platelet-derived growth factor (PDGF): amino-terminal amino
 RT acid sequence.";
 RL Science 220:963-965(1983).
 RN [12]
 RP SEQUENCE OF 82-112.
 RX MEDLINE; 83244981.
 RA Waterfield M.D., Strace G.T., Whittle N., Stroobant P., Johansson A.,
 RA Westesson A., Westermark B., Heldin C.H., Huang J.S., Deuel T.F.;
 RT "Platelet-derived growth factor is structurally related to the
 RL putative transforming protein p28sis of simian sarcoma virus.";
 RN Nature 304:35-39(1983).
 RN [13]
 RP MUTAGENESIS, & IMPORTANCE OF ARG-108 AND ILE-111 FOR RECEPTOR-BINDING.
 RX MEDLINE; 92097530.
 RA Clements J.M., Bawden L.J., Bloxidge R.E., Catlin G., Cook A.L.,
 RA Craig S., Drummond A.H., Edwards R.M., Fallon A., Green D.R.,
 RA Hellewell P.G., Kirwin P.M., Nacey P.D., Richardson S.J., Brown D.,
 RA Chahwa S.B., Snarey M., Winslow D.;
 RT "Two PDGF-B chain residues, arginine 27 and isoleucine 30, mediate
 RL receptor binding and activation.";
 RN EMBO J. 10:4113-4120(1991).
 RN [14]
 RP INTERCHAIN DISULFIDE BONDS.
 RX MEDLINE; 92283833.
 RA Andersson M., Oestman A., Baekstroem G., Hellman U.,
 RA George-Nascimento C., Westermark B., Heldin C.-H.;
 RT "Assignment of interchain disulfide bonds in platelet-derived growth
 RL factor (PDGF) and evidence for agonist activity of monomeric PDGF.";
 RN J. Biol. Chem. 267:11260-11266(1992).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE; 93010987.
 RA Osner C., D'Arcy A., Winkler F.K., Eggmann B., Hosang M.;
 RT "Crystal structure of human platelet-derived growth factor BB.";
 RL EMBO J. 11:3921-3926(1992).
 CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
 CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
 CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
 CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
 CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
 CC -1- SUBUNIT: ANTI-PAALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
 CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
 CC TRANSFORMATION PROCESSES.
 CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAME REGRANEX (ORTHO-MCNEIL).
 CC USED TO PROMOTE HEALING IN DIABETIC NEUROPATHIC FOOT ULCERS.
 CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
 CC PDGF RECEPTOR.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -1- DATABASE: NAME-RED systems' cytokine source book;
 CC WWW="http://www.rndsystems.com/cyt_cat/pdgrf.html".
 CC -1- DATABASE: NAME-Regranex; NOTE-Clinical information on Regranex;
 CC WWW="http://www.regranex.com/".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; K01401; AAA60552.1; -
 CC EMBL; K01918; AAA60552.1; JOINED.
 CC EMBL; J00121; AAA60552.1; JOINED.
 CC EMBL; K01398; AAA60552.1; JOINED.
 CC EMBL; K01399; AAA60552.1; JOINED.
 CC EMBL; K01400; AAA60552.1; JOINED.
 CC EMBL; K02811; CAA26579.1; -
 CC EMBL; M12783; AAA60553.1; -
 CC EMBL; X02744; CAA26524.1; -
 CC EMBL; K01917; AAA98793.1; -
 CC EMBL; K01913; AAA98793.1; JOINED.
 CC EMBL; K01914; AAA98793.1; JOINED.
 CC EMBL; K01915; AAA98793.1; JOINED.
 CC EMBL; K01916; AAA98793.1; JOINED.
 CC EMBL; X03702; CAA27333.1; -
 CC EMBL; Z81010; CAB02635.1; -
 CC EMBL; X00561; CAA25228.1; -
 CC EMBL; X00561; CAA25229.1; -
 CC EMBL; X98706; CAA67262.1; -
 CC PIR; A94276; PFHUG2.
 CC PDB; LPDG; 31-JAN-94.
 CC MIM; 190040; -
 CC PFAM; PF00341; PDGF; 1.
 CC PRINTS; PR00438; GFCYSKNOT.
 CC PROSITE; PS00249; PDGF_1; 1.
 CC KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal;
 KW Pharmaceutical; 3D-structure.
 FT SIGNAL 1 20
 FT PROPEP 21 81
 FT CHAIN 82 190
 FT CHAIN 191 241
 FT PROPEP 108 108
 FT SITE 111 111
 FT SITE 111 111
 FT DISULFID 97 141
 FT DISULFID 130 178
 FT DISULFID 134 180
 FT DISULFID 124 124
 FT DISULFID 133 133
 FT DISULFID 133 133
 FT CONFLICT 21 21
 FT CONFLICT 101 101
 FT CONFLICT 105 105
 FT CONFLICT 107 107
 FT STRAND 90 91
 FT STRAND 94 94
 FT STRAND 98 103
 FT STRAND 120 121
 FT STRAND 124 131
 FT TURN 137 138
 CC ...
 CC Note: remainder of annotations omitted.
 CC Query Match 5.98; Score 6; DB 1; Length 241;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC Db 95 AECKTR 100
 CC QY 71 AECKTR 76
 CC RESULT 32
 CC ID PDGF_FELCA STANDARD; PRT; 245 AA.
 CC AC P12919;
 CC DT 01-OCT-1989 (Rel. 12, Created)
 CC DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
DE (PDGFB) (PDGFB-2).
GN PDGFB OR C-SIS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87146463.
RA van den Ouweland A.M.W., van Groningen J.J.M., Schalken J.A.,
RA van Neck H.W., Bloemers H.P.J., van de Ven W.J.M.;
RT "Genetic organization of the c-sis transcription unit.";
RL Nucleic Acids Res. 15:959-970(1987).
RN [2]
RP REVISIONS.
RA van den Ouweland A.M.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC -!- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC -!- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X05112; CAA28758.1; ALT_SEQ.
CC PIR; A26402; TVCTSS.
CC HSSP; P01127; LPDG.
CC PFAM; PF00341; PDGF_1.
CC PRINTS; PR00438; GFCYSKNOT.
CC DR PROSITE; PS00249; PDGF_1; 1.
CC KW Mitogen; Growth factor; Proto-oncogene; Platelet; Signal.
CC FT SIGNAL 1 20
CC FT PROPEP 21 81
CC FT CHAIN 82 194
CC FT PROPEP 195 245
CC FT DISULFID 101 145
CC FT DISULFID 134 182
CC FT DISULFID 138 184
CC FT DISULFID 128 128
CC FT DISULFID 137 137
CC FT DISULFID 137 137
CC SEQUENCE 245 AA; 27787 MW; E771529ID9837512 CRC64;

CC Query Match 5.9%; Score 6; DB 1; Length 245;
CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 AECCTR 104
QY 71 AECCTR 76

RESULT 33
ID VGL2_EBV STANDARD; PRT; 248 AA.
AC P03218;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE PROBABLE MEMBRANE GLYCOPROTEIN.
GN BILF2.

OS Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85035713.
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RT Epstein-Barr virus.;
RL Mol. Biol. Med. 1:21-45(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84270667.
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90244363.
RA Mackett M., Conway M.J., Arrand J.R., Haddad R.S.,
RA Hutt-Fletcher L.M.;
RT "Characterization and expression of a glycoprotein encoded by the
RT Epstein-Barr virus BamHI I fragment.";
RL J. Virol. 64:2545-2552(1990).
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01555; CAA24803.1; -.
CC DR EMBL; M37129; AAA45876.1; -.
CC PIR; A03780; QQBE4L.
CC DR PIR; S33052; S33052.
CC DR PFAM; PF00047; Ig_1.
CC KW Glycoprotein; Late protein; Membrane.
CC SEQUENCE 248 AA; 27076 MW; C3F33A253B959ADA CRC64;

CC Query Match 5.9%; Score 6; DB 1; Length 248;
CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 TNGGYN 83
QY 25 TNGGYN 30

RESULT 34
ID YKT0_CAEEL STANDARD; PRT; 254 AA.
AC P34321;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HYPOTHETICAL 29.1 KDA PROTEIN C07A9.10 IN CHROMOSOME III.
GN C07A9.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: TO C.ELEGANS F52C9.6 AND F23C11.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29094; CAA82337.1; -;
CC PIR; S40702; S40702.
CC WORMPEP; C07A9.10; CE00496.
CC PFAM; PF00442; UCH-1; 1.
CC Hypothetical protein.
SQ SEQUENCE 254 AA; 29082 MW; 8057BF17FE13CB50 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 186 LAAECK 191
| | | | |
Qy 69 LAAECK 74
RESULT 35
ID RNPLAQUE STANDARD; PRT; 255 AA.
AC 067069;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-FEB-2000 (Rel. 39, Last annotation update)
DE RIBONUCLEASE PH (EC 2.7.7.56) (RNASE PH) (TRNA
DN NUCLEOTIDYLTRANSFERASE).
GN RPH OR RNPH OR AQ-924.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-VF5;
CC MEDLINE; 98196666.
CC Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
CC "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus".
CC Nature 332:353-358(1998).
CC -1- FUNCTION: RNASE PH IS A PHOSPHORYLATED EXORIBONUCLEASE THAT
CC REMOVES NUCLEOTIDE RESIDUES FOLLOWING THE -CCA TERMINUS OF TRNA
CC AND ADDS NUCLEOTIDES TO THE ENDS OF RNA MOLECULES BY USING
CC NUCLEOSIDE DIPHOSPHATES AS SUBSTRATES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: TRNA(N+1) + ORTHOPHOSPHATE = TRNA(N) +
CC A NUCLEOSIDE DIPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000714; AAC07032.1; -;
CC PFAM; PF01138; RNase_PH; 1.
DR

DR PROSITE; PS01277; RIBONUCLEASE_PH; 1.
KW Transferase; Nucleotidyltransferase; tRNA processing.
SQ SEQUENCE 255 AA; 28372 MW; 7131CCD8AEDC3917 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 42 SVIENV 47
| | | | |
Qy 38 SVIENV 43
RESULT 36
ID MS2A_DROME STANDARD; PRT; 264 AA.
AC P1033;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MALE ACCESSORY GLAND SECRETORY PROTEIN 355A PRECURSOR.
GN ACP26AA OR MST26AA OR MST355A.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-CANTON-S;
CC MEDLINE; 89053045.
CC Monsma S.A., Wolfner M.F.;
CC "Structure and expression of a Drosophila male accessory gland gene
CC RT whose product resembles a peptide pheromone precursor.";
CC RL Genes Dev. 2:1063-1073(1988).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN-NC1, NC2, NC3, NC4, NC5, NC6, NC7, NC8, NC9 AND NC10;
CC MEDLINE; 93106377.
CC Aguade M., Miyashita N., Langley C.H.;
CC "Polymorphism and divergence in the Mst26A male accessory gland gene
CC region in Drosophila".
CC RL Genetics 132:755-770(1992).
CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE'S
CC HEMOLymph DURING MATING, AFFECTING EGG-LAYING AND BEHAVIOR AFTER
CC MATING.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES.
CC -1- PTM: IT UNDERGOES SEVERAL CLEAVAGES AS IT IS SECRETED AND IT
CC IS FURTHER PROCESSED IN THE RECIPIENT FEMALE.
CC -1- SIMILARITY: REGION OF HOMOLOGY WITH APLYSIA CALIFORNICA EGG-LAYING
CC HORMONE (ELH).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00219; CAA69366.1; -;
CC EMBL; X70888; CAA50232.1; -;
CC EMBL; X70889; CAA50234.1; -;
CC EMBL; X70890; CAA50236.1; -;
CC EMBL; X70891; CAA50238.1; -;
CC EMBL; X70892; CAA50240.1; -;
CC EMBL; X70893; CAA50242.1; -;
CC EMBL; X70894; CAA50244.1; -;
CC EMBL; X70895; CAA50246.1; -;
CC EMBL; X70896; CAA50248.1; -;
CC EMBL; X70897; E64715; ALT_FRAME.
CC PIR; S02853; S02853.
CC PIR; S30409; S30409.
CC PIR; S30411; S30411.
DR

DR PIR: S30413; S30413.
DR PIR: S30415; S30415.
DR PIR: S30423; S30423.
DR PIR: S30425; S30425.
DR FLYBASE; FBgn002855; Acp26Aa.
KW Glycoprotein; Behavior; Signal.
FT SIGNAL 1 18
FT CHAIN 19 264
FT
FT CARBOHYD 88
FT FT POTENTIAL.
FT FT 355A.
FT FT CARBOHYD 122 122
FT FT CARBOHYD 138 138
FT FT CARBOHYD 145 145
FT FT CARBOHYD 145 145
FT FT POTENTIAL.
FT FT TO APLYSIA CALIFORNICA ELH.
FT FT N -> S (IN STRAINS NC1, NC2, NC3, NC4,
FT FT NC6, NC8, NC9).
FT FT Q -> K (IN STRAINS NC5, NC7, NC10).
FT FT L -> Q (IN STRAIN NC9).
FT FT P -> T (IN STRAINS NC2, NC3, NC9).
FT FT D -> N (IN STRAIN NC3).
FT FT D -> N (IN STRAINS NC4, NC6, NC8).
FT FT N -> S (IN STRAINS NC1, NC2, NC3, NC9).
FT FT L -> I (IN STRAINS NC1, NC2, NC4, NC6,
FT FT NC8, NC9).
FT FT S -> I (IN STRAIN NC2).
FT FT R -> K (IN STRAINS NC2, NC11).
FT FT
SQ SEQUENCE 264 AA; 29671 MW; BD64298CE04DFDE5 CRC64;
Query Match 5.9%; Score 6; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 96 SSIDLN 101
Yy 32 SSIDLN 37
RESULT 37-
ID YAIM_HAEIN STANDARD; PRT; 275 AA.
AC P4456;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHEICAL PROTEIN HI1084.
GN HI0184.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
CC Haemophilus.
CC [1]
RP SEQUENCE FROM N.A.
RX STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McClellan J., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- SIMILARITY: STRONG, TO E.COLI YAIM AND YEIG.
CC -!- SIMILARITY: STRONG, TO HUMAN ESTERASE D.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

CC or send an email to license@isb-sib.ch.
CC
CC EMBL; U32703; AAC21853.1; -.
DR TIGR; HI0184;
DR PFAM; PF00756; Esterase; 1.
KW Hypothetical protein.
SQ SEQUENCE 275 AA; 31317 MW; 781F5C0411546D3D CRC64;
Query Match 5.9%; Score 6; DB 1; Length 275;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 233 FIETCR 238
Yy 54 FIETCR 59
RESULT 38
ID HEM3_CHLVI STANDARD; PRT; 279 AA.
AC P28464;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PORPHOBILIN GEN DEAMINASE (EC 4.3.1.8) (PBG) (HYDROXYMETHYLBILANE
DE SYNTHASE) (HMBS) (PRE-UROPORPHYRINOGEN SYNTHASE).
GN HEMC.
OS Chlorobium vibrioforme.
OC Bacteria; Green sulfur bacteria; Chlorobium.
CC [1]
RP SEQUENCE FROM N.A.
RX STRAIN-F. THIOSULFATOPHILUM / NCIB 8327;
RX MEDLINE; 92171712.
RA Majumdar D., Avissar Y.J., Wyche J.H., Beale S.I.;
RT "Structure and expression of the Chlorobium vibrioforme hema gene.";
RL Arch. Microbiol. 156:281-289(1991).
CC -!- FUNCTION: TETRAPOLYMERIZATION OF THE MONOPYRROLE PBG INTO THE
CC HYDROXYMETHYLBILANE PREUROPORPHYRINOGEN IN SEVERAL DISCRETE STEPS.
CC -!- CATALYTIC ACTIVITY: 4 PORPHOBILINOGEN + H(2)O =
CC HYDROXYMETHYLBILANE + 4 NH(3).
CC -!- COFACTOR: COVALENTLY BINDS A DIPYRROMETHANE COFACTOR TO WHICH
CC THE PORPHOBILINOGEN SUBUNITS ARE ADDED.
CC -!- PATHWAY: FOURTH STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC INVOLVED IN CHLOROPHYLL BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE HMBS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

Query Match 5.9%; Score 6; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 IDGTLK 224

|||||

Qy 94 IDGTLK 99

RESULT 39

ID YLMW9CAEEL STANDARD; PRT; 288 AA.
AC P34412; 1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL 32.9 KDA PROTEIN F22B7.9 IN CHROMOSOME III.
GN F22B7.9
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RM MEDLINE; 94150718.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L12018; AAA65465.1; -

DR WORKPEP; F22B7.9; CE00162.

KW Hypothetical protein.

SEQUENCE 288 AA; 32856 MW; 2E7583B33DD23F8C CRC64;

Query Match 5.9%; Score 6; DB 1; Length 288;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 199 NVDGSL 204

|||||

Qy 42 NVDGSL 47

RESULT 40

ID PYR3_PREDI STANDARD; PRT; 288 AA.
AC P11400;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE PHYCOBILISOME 39 KDA LINKER POLYPEPTIDE, PHYCOCYANIN-ASSOCIATED, ROD
DE (L-39/R).
GN CPC12.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
[1]
RN SEQUENCE FROM N.A.
RP SEQUENCE 288 AA; 32856 MW; 2E7583B33DD23F8C CRC64;
RM MEDLINE; 87222193.

RA Lomax T.L., Conley P.B., Schilling J., Grossman A.R.;
RT "Isolation and characterization of light-regulated phycobilisome
RT linker polypeptide genes and their transcription as a polycistronic
RT mRNA";
RL J. Bacteriol. 169:2675-2684(1987).
CC -1- FUNCTION: ROD LINKER PROTEIN, ASSOCIATED WITH PHYCOCYANIN. LINKER
CC POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION AND THE LOCATION
CC OF THE DISC-SHAPED PHYCOBILIPROTEIN UNITS WITHIN THE PHYCOBILISOME
CC AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN ORDER TO MEDIATE A
CC DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN OCCURS IN THE ROD, IT IS
CC ASSOCIATED WITH PHYCOCYANIN.
CC -1- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M16490; AAA24887.1; -

DR PIR; B25974; B25974.

DR PFAM; PF01383; CpcB; 1.

DR PFAM; PF00427; PBS_Linker_poly; 1.

KW Phycobilisome; Photosynthesis.

FT INIT_MET 0

SEQUENCE 288 AA; 32184 MW; D5A325E066A0B2F6 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 288;

Best Local Similarity 100.0%; Pred. No. 8.75e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 QLAGSS 186

|||||

Qy 62 QLAGSS 67

RESULT 41

ID RS2_RICPR STANDARD; PRT; 296 AA.
AC Q3ZE61;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S2.
GN RPSB OR RP086.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE; 99039499.
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AJ235270; CAAL4556.1; -
DR PFAM; PF00318; Ribosomal_S2; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.

```

DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein.
SQ SEQUENCE 296 AA; 32961 MW; 386B925D43BD6C0 CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 72 FVSTKI 77
    |||||
Qy 80 FVSTKI 85

RESULT 42
ID SFMH_ECOLI STANDARD; PRT; 325 AA.
AC P75715; P77078;
DT 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SFMH PROTEIN PRECURSOR.
GK SFMH.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Nanath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN REGULATION OF LENGTH AND MEDIATION OF
CC ADHESION OF SFMA FIMBRIAE (BUT NOT NECESSARY FOR THE PRODUCTION
CC OF FIMBRIAE). SEEMS TO BE MANNOSE BINDING ADHESIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FIMH / LPFD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000159; AAC73635.1; -.
CC EMBL; U82598; AAB40731.1; ALT_INIT.
CC ECOGENE; EGI3884; SFMH.
KW Fimbria; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? 325 SFMH PROTEIN.
SQ SEQUENCE 325 AA; 35493 MW; B607915A95542193 CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 278 DLNSVI 283
    |||||
Qy 35 DLNSVI 40

RESULT 43
ID RLUD_BORBU STANDARD; PRT; 326 AA.
AC P70870; O51050;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE D (EC 4.2.1.70)
DE (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLASE).
GN RLUD OR BB0018.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB19;
RA Boursaux-Eude C., Margarita D., Belfaiza J., Old I.G.,
RA Saint-Girons I.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: RESPONSIBLE FOR SYNTHESIS OF PSEUDOURIDINE FROM URACIL
CC AT TWO POSITIONS IN 23S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: URACIL + D-RIBOSE 5-PHOSPHATE = PSEUDOURIDINE
CC 5'-PHOSPHATE + H(2)O.
CC -!- SIMILARITY: BELONGS TO THE RLU FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y09142; CAA70352.1; -.
CC EMBL; AE001116; AAC66413.1; ALT_INIT.
DR TIGR; BB0018; -.
DR PFAM; PF00849; Pseudou synth_2; 1.
DR PROSITE; PS01129; PSI_RLU; 1.
KW Lyase.
FT CONFLICT 172 172 T -> I (IN REF. 1).
FT CONFLICT 268 268 K -> E (IN REF. 1).
FT CONFLICT 299 326 FVFLSNFYKSDLNLIIDNLVFLRDF ->
FT ICYFFVXFLQE (IN REF. 1).
SQ SEQUENCE 326 AA; 38067 MW; 52266C0AF8BEA3C CRC64;

Query Match
  5.9%; Score 6; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 263 VDGLSK 268
    |||||
Qy 43 VDGLSK 48

RESULT 44
ID VGH_BPPHX STANDARD; PRT; 328 AA.
AC P03646;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
GN H.
OS Bacteriophage phi-X174.

```

OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 77171175.
 RA Sanger F., Air G.M., Barrell B.G., Brown N.L., Coulson A.R.,
 RA Frides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RA "Nucleotide sequence of bacteriophage phi x174 DNA.";
 RL Nature 265:687-695(1977).
 RN [2]
 RP REVISIONS.
 RX MEDLINE; 79091185.
 RA Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
 RA Brown N.L., Frides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RA "The nucleotide sequence of bacteriophage phi x174.";
 RL J. Mol. Biol. 125:225-246(1978).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE; 92107297.
 RA McKenna R., Xia D., Williangmann P., Ilag L.L., Krishnaswamy S.,
 RA Rossmann M.G., Olson N.H., Baker T.S., Incardona N.L.;
 RA "Atomic structure of single-stranded DNA bacteriophage phi x174 and
 RA its functional implications.";
 RL Nature 355:137-143(1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE; 94210479.
 RA McKenna R., Ilag L.L., Rossmann M.G.;
 RA "Analysis of the single-stranded DNA bacteriophage phi x174, refined
 RA at a resolution of 3.0 A";
 RL J. Mol. Biol. 237:517-543(1994).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 1130:277-288(1992).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92223109.
 RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
 RA "Nucleotide sequence of the genome of the bacteriophage alpha 3:
 RA interrelationship of the genome structure and the gene products with
 RA those of the phages, phi x174, G4 and phi K.";
 RL Blochim. Biophys. Acta 1130:277-288(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 85226468.
 RA Kodaira K.-I., Nakano K., Taketo A.;
 RA "Function and structure of microvirid phage alpha 3 genome. DNA
 RA sequence of H gene and properties of missense H mutant.";
 RL Blochim. Biophys. Acta 825:255-260(1985).
 RN [3]
 RP SEQUENCE OF 1-70 FROM N.A.
 RX MEDLINE; 80049950.
 RA Sims J., Capon D., Dressler D.;
 RA "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RA sequences of the negative strand initiation sites of bacteriophages
 RA St-1, phi K, and alpha 3.";
 RL J. Biol. Chem. 254:12615-12628(1979).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 825:255-260(1985).
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86165869.
 RA Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
 RA Brown N.L., Frides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RA "The nucleotide sequence of bacteriophage phi x174 DNA.";
 RL Nature 265:687-695(1977).
 RN [2]
 RP REVISIONS.
 RX MEDLINE; 79091185.
 RA Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
 RA Brown N.L., Frides J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RA "The nucleotide sequence of bacteriophage phi x174.";
 RL J. Mol. Biol. 125:225-246(1978).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE; 92107297.
 RA McKenna R., Xia D., Williangmann P., Ilag L.L., Krishnaswamy S.,
 RA Rossmann M.G., Olson N.H., Baker T.S., Incardona N.L.;
 RA "Atomic structure of single-stranded DNA bacteriophage phi x174 and
 RA its functional implications.";
 RL Nature 355:137-143(1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE; 94210479.
 RA McKenna R., Ilag L.L., Rossmann M.G.;
 RA "Analysis of the single-stranded DNA bacteriophage phi x174, refined
 RA at a resolution of 3.0 A";
 RL J. Mol. Biol. 237:517-543(1994).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 1130:277-288(1992).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92223109.
 RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
 RA "Nucleotide sequence of the genome of the bacteriophage alpha 3:
 RA interrelationship of the genome structure and the gene products with
 RA those of the phages, phi x174, G4 and phi K.";
 RL Blochim. Biophys. Acta 1130:277-288(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 85226468.
 RA Kodaira K.-I., Nakano K., Taketo A.;
 RA "Function and structure of microvirid phage alpha 3 genome. DNA
 RA sequence of H gene and properties of missense H mutant.";
 RL Blochim. Biophys. Acta 825:255-260(1985).
 RN [3]
 RP SEQUENCE OF 1-70 FROM N.A.
 RX MEDLINE; 80049950.
 RA Sims J., Capon D., Dressler D.;
 RA "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RA sequences of the negative strand initiation sites of bacteriophages
 RA St-1, phi K, and alpha 3.";
 RL J. Biol. Chem. 254:12615-12628(1979).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 825:255-260(1985).
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86165869.

RA Lau P.C.K., Spencer J.H.;
 RT "Nucleotide sequence and genome organization of bacteriophage S13
 RT DNA.";
 RL Gene 40:273-284(1985).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 825:255-260(1985).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 80049950.
 RA Sims J., Capon D., Dressler D.;
 RA "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RA sequences of the negative strand initiation sites of bacteriophages
 RA St-1, phi K, and alpha 3.";
 RL J. Biol. Chem. 254:12615-12628(1979).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 825:255-260(1985).
 CC [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 85226468.
 RA Kodaira K.-I., Nakano K., Taketo A.;
 RA "Function and structure of microvirid phage alpha 3 genome. DNA
 RA sequence of H gene and properties of missense H mutant.";
 RL Blochim. Biophys. Acta 825:255-260(1985).
 RN [3]
 RP SEQUENCE OF 1-70 FROM N.A.
 RX MEDLINE; 80049950.
 RA Sims J., Capon D., Dressler D.;
 RA "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RA sequences of the negative strand initiation sites of bacteriophages
 RA St-1, phi K, and alpha 3.";
 RL J. Biol. Chem. 254:12615-12628(1979).
 CC -1- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INTERRELATIONSHIP OF THE GENOME STRUCTURE AND THE GENE PRODUCTS WITH
 CC THOSE OF THE PHAGES, PHI X174, G4 AND PHI K.;
 CC BLOCHIM. BIOPHYS. ACTA 825:255-260(1985).
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86165869.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X60322; CAA42883.1; -
 DR EMBL; M25640; AAA32174.1; -
 DR EMBL; J02444; AAA32176.1; -
 DR PIR; A04257; ZHBPA3.
 DR PIR; A21537; A21537.
 DR PIR; S22332; S22332.
 KW Coat protein.
 SQ SEQUENCE 330 AA; 34844 MW; B194EF4A83CE3CE8 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 58 SAIQGS 63
 |||||
 Qy 11 SAIQGS 16

RESULT 47
 ID VGH_BPPHK STANDARD; PRT; 332 AA.

AC P03649;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
 GN H.
 OS Bacteriophage phi-K.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-42 FROM N.A.
 RX MEDLINE; 80049950.

RA Sims J., Capon D., Dressler D.;
 RT "dnaG (primase)-dependent origins of DNA replication. Nucleotide
 RT sequences of the negative strand initiation sites of bacteriophages
 RT St-1, phi K, and alpha 3".
 RT J. Biol. Chem. 254:12613-12628(1979).
 CC -!- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
 CC -!- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES OF THE F, G, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
 CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X60323; CAA42893.1; -
 DR EMBL; M10726; AAA32365.1; -
 DR PIR; B04256; B04256.
 KW Coat protein.
 SQ SEQUENCE 332 AA; 35110 MW; C8959595D1F9C5A7A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 332;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 60 SAIQGS 65
 |||||
 Qy 11 SAIQGS 16

RESULT 48
 ID VGH_BPG4 STANDARD; PRT; 337 AA.

AC P03647;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MINOR SPIKE PROTEIN (H PROTEIN) (PILOT PROTEIN).
 GN H.
 OS Bacteriophage G4.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 79053264.
 RA Godson G.N., Barrell B.G., Staden R., Fiddes J.C.;
 RT "Nucleotide sequence of bacteriophage G4 DNA".
 RL Nature 276:236-247(1978).
 CC -!- FUNCTION: MINOR SPIKE COMPONENT OF THE VIRAL SHELL. H PROTEIN IS
 CC INVOLVED IN THE EJECTION OF THE PHAGE DNA IN THE HOST AND IS
 CC INJECTED WITH THE DNA IN THE PERIPLASMIC SPACE OF THE HOST.
 CC -!- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
 CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J02454; AAA32325.1; -
 DR EMBL; V00657; CAA24021.1; -
 DR PIR; A04255; ZHBPG4.
 KW Coat protein.
 SQ SEQUENCE 337 AA; 35706 MW; 3D8E94F2E21CB9AB CRC64;

Query Match 5.9%; Score 6; DB 1; Length 337;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 SAIQGS 59
 |||||
 Qy 11 SAIQGS 16

RESULT 49
 ID YG5F_YEAST STANDARD; PRT; 366 AA.

AC P53320;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE MITOCHONDRIAL CARRIER YGR257C.
 GN YGR257C OR G9175.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE; 97279233.
 RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F.,
 RA Frontali L.;
 RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast
 RT chromosome VII reveals the presence of three new open reading frames
 RT and of a tRNA^{Thr} gene".
 RL Yeast 13:369-372(1997).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL

CC INNER MEMBRANE (POTENTIAL).
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X99228; CAA67613.1; -;
 CC EMBL; Z73042; CAA97286.1; -;
 CC PFAM; PF00153; mito_carr; 2;
 CC PROSITE; PS00215; MITOCH_CARRIER; 1;
 CC KW Hypothetical protein: Mitochondrion; Inner membrane; Repeat;
 CC Transmembrane; Transport.
 CC TRANSMEM 17 36 POTENTIAL.
 CC TRANSMEM 124 140 POTENTIAL.
 CC TRANSMEM 162 182 POTENTIAL.
 CC TRANSMEM 268 286 POTENTIAL.
 CC SEQUENCE 366 AA; 40763 MW; B0358B6E818CB1E CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 366;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 22 GSVLTS 27
 CC Y 15 GSVLTS 20
 CC
 CC RESULT 50
 CC ID AARD_HUMAN STANDARD; PRT; 398 AA.
 CC AC P22760;
 CC DT 01-AUG-1991 (Rel. 19, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE ARYLACETAMIDE DEACETYLASE (EC 3.1.1.-) (AADAC).
 CC GN DAC.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC [1]
 CC RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC RP TISSUE=LIVER;
 CC RX MEDLINE; 94342357.
 CC RA Probst M.R., Beer M., Beer D., Jenoe P., Meyer U.A., Gasser R.;
 CC "Human liver arylacetamide deacetylase. Molecular cloning of a novel
 CC esterase involved in the metabolic activation of arylamine
 CC carcinogens with high sequence similarity to hormone-sensitive
 CC lipase.";
 CC RL J Biol. Chem. 269:21650-21656(1994).
 CC [2]
 CC RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
 CC TX TISSUE=LIVER;
 CC RX MEDLINE; 91254316.
 CC RA Probst M.R., Jenoe P., Meyer U.A.;
 CC "Purification and characterization of a human liver arylacetamide
 CC deacetylase.";
 CC RL Biochem. Biophys. Res. Commun. 177:453-459(1991).
 CC [1-] FUNCTION: ARYLACETAMIDE DEACETYLATION IS AN IMPORTANT ENZYME
 CC ACTIVITY IN THE METABOLIC ACTIVATION OF ARYLAMINE SUBSTRATES TO
 CC ULTIMATE CARCINOGENS.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MICROSOMAL
 CC (POTENTIAL).
 CC -!- TISSUE SPECIFICITY: LIVER.
 CC -!- SIMILARITY: BELONGS TO THE "GDXG" FAMILY OF LIPOLYTIC ENZYMES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L32179; AAA35551.1; -;
 CC MIN; G00338; -;
 CC PFAM; PF00135; Coesterase; 1;
 CC PROSITE; PS01174; LIPASE_GDXG_SER; 1;
 CC KW Hydrolase; Transmembrane; Microsome; Signal-anchor.
 CC INIT_MET 0 0
 CC TRANSMEM 4 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC ACT_SITE 110 110 (POTENTIAL).
 CC ACT_SITE 188 188 POTENTIAL.
 CC FT CONFLICT 2 2 R -> M (IN REF. 2).
 CC SQ SEQUENCE 398 AA; 45540 MW; 201FD6AC1700AD0C CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 398;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 302 GSSELA 307
 CC Y 65 GSSELA 70
 CC
 CC RESULT 51
 CC ID PAP3_XENLA STANDARD; PRT; 400 AA.
 CC AC P51006;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
 CC DE POLY(A) POLYMERASE TYPE 3 (EC 2.7.7.19) (PAP) (POLYNUCLEOTIDE
 CC DE ADENYLTRANSFERASE) (FRAGMENT).
 CC OS Xenopus laevis (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 CC OC Xenopodinae; Xenopus.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX TISSUE=OVARY;
 CC RA MEDLINE; 96079940.
 CC RA Ballantyne S., Bilger A., Astrom J., Virtanen A., Wickens M.;
 CC "Poly (A) polymerases in the nucleus and cytoplasm of frog oocytes:
 CC dynamic changes during oocyte maturation and early development.";
 CC RL RNA 1:64-78(1995).
 CC [1-] FUNCTION: POLYMERASE THAT CREATES THE 3' POLY(A) TAIL OF MRNA'S.
 CC MAY ACQUIRE SPECIFICITY THROUGH INTERACTION WITH A CLEAVAGE AND
 CC POLYADENYLATION FACTOR (CPSF).
 CC -!- CATALYTIC ACTIVITY: N ATP + (NUCLEOTIDE)(M) - N PYROPHOSPHATE +
 CC (NUCLEOTIDE)(M+N).
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U19975; AAC59747.1; -;
 CC PFAM; PF01909; NTP_transf_2; 1;
 CC KW mRNA processing; Transferase; Transcription; RNA-binding;
 CC Nucleic acid protein.
 CC ACT_SITE 110 110 BY SIMILARITY.
 CC ACT_SITE 112 112 BY SIMILARITY.
 CC ACT_SITE 164 164 BY SIMILARITY.
 CC FT DOMAIN 382 390 NUCLEAR LOCALIZATION SIGNAL 1 (BY
 CC FT NON_TER 400 400 SIMILARITY).
 CC

SQ SEQUENCE 400 AA; 46020 MW; AB4347C074E50CA3 CRC64;
 Query Match 5.9%; Score 6; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 85 SVIENV 90
 |||||
 Qy 38 SVIENV 43

RESULT 52
 ID CAG5_CHICK STANDARD; PRT; 404 AA.
 AC Q92184;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE ALPHA-N-ACETYL GALACTOSAMINIDE ALPHA-2,6-SIALYLTRANSFERASE
 (EC 2.4.99.-) (SF6GALNACII).
 Gallus gallus (Chicken).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RX MEDLINE; 94308168.
 RA Kurosawa N., Kojima N., Inoue M., Hamamoto T., Tsuji S.;
 RT "Cloning and expression of Gal beta 1,3galnac-specific GalNAc alpha
 2,6-sialyltransferase.";
 RT J. Biol. Chem. 269:19048-19053(1994).
 RL CC -1- CATALYTIC ACTIVITY: CMP-N-ACETYLNEURAMINATE + GLYCANO-BETA-D-
 GALACTOSYL-1,3-(N-ACETYL-D-GALACTOSAMINYL)-GLYCOPROTEIN = CMP
 + GLYCANO-BETA-D-GALACTOSYL-(2,6-ALPHA-N-ACETYLNEURAMINYL)-
 (N-ACETYL-D-GALACTOSAMINYL)-GLYCOPROTEIN.
 CC -1- PATHWAY: GLYCOSYLATION.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: HEART, KIDNEY, TESTES, BRAIN, LIVER AND LUNG.
 CC -1- DEVELOPMENTAL STAGE: ABUNDANTLY EXPRESSED AT ALL EMBRYONIC STAGES
 BUT NOT PRESENT IN ADULT TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE VERTEBRATE SIALYLTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X77775; CAA54813.1; -;
 DR PFAM; PF00777; Sialyltransf. 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack.
 FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 9 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL).
 FT DOMAIN 26 404 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 161 161 POTENTIAL.
 FT CARBOHYD 191 191 POTENTIAL.
 SQ SEQUENCE 404 AA; 45826 MW; DCC177AA01ABB60A CRC64;

Query Match 5.9%; Score 6; DB 1; Length 404;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 272 SAIQGS 277
 |||||
 Qy 11 SAIQGS 16

RESULT 53
 ID DEOB_ECOLI STANDARD; PRT; 407 AA.

P07651;
 AC 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PHOSPHOTENOMUTASE (EC 5.4.2.7) (PHOSPHODEOXYRIBOMUTASE).
 GN DEOB OR DRM OR THYR.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE; 95334362.
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes.";
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN [2]
 RP SEQUENCE OF 1-37 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE; 84272212.
 RA Valentin-Hansen P., Hammer K., Larsen J.E.L., Svendsen I.;
 RT "The internal regulated promoter of the deo operon of Escherichia
 coli K-12.";
 RL Nucleic Acids Res. 12:5211-5224(1984).
 CC -1- CATALYTIC ACTIVITY: D-RIBOSE 1-PHOSPHATE - D-RIBOSE 5-PHOSPHATE.
 CC -1- CATALYTIC ACTIVITY: 2-DEOXY-D-RIBOSE 1-PHOSPHATE - 2-DEOXY-
 D-RIBOSE 5-PHOSPHATE.
 CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOTENOMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U14003; AAA97279.1; -;
 DR EMBL; AE000508; AAC77336.1; -;
 DR EMBL; X00742; CAA25325.1; -;
 DR PIR; B22909; B22909.
 DR ECOGENE; EG10220; DEOB.
 DR PFAM; PF01676; Metalloenzyme; 1.
 KW Isomerase.
 SQ SEQUENCE 407 AA; 44370 MW; 516F3018DC77A077 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 407;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 202 TNGYV 207
 |||||
 Qy 25 TNGYV 30

RESULT 54
 ID GLYA_BACSU STANDARD; PRT; 415 AA.
 AC P39148;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
 (SHMT).
 DE (SHMT).
 GN GLYA OR GLYC OR IPC-34D.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;

RA Glaser P., Danchin A.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 238002; CAA86110.1; -;
 CC EMBL; 299122; CAB15707.1; -;
 CC SUBTILIST; BG10944; GLYA.
 CC PFAM; PF00464; SHMT; 1.
 CC PROSITE; PS00096; SHMT; 1.
 CC Transferase; Pyridoxal phosphate: One-carbon metabolism.
 CC FT BINDING 226 226 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC SEQUENCE 415 AA; 45489 MW; A3AD2F7C40AB14A1 CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 415;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 43 QGSVLT 48
 QY |||||
 14 QGSVLT 19

 CC RESULT 55
 CC ID GLYA_STRCO STANDARD; PRT; 420 AA.
 CC AC O86565;
 CC DT 15-FEB-2000 (Rel. 39, Created)
 CC DT 15-FEB-2000 (Rel. 39, Last sequence update)
 CC DT 15-FEB-2000 (Rel. 39, Last annotation update)
 CC DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
 CC (SHMT).
 CC GN GLYA OR SC2A11.04C.
 CC OS Streptomyces coelicolor.
 CC OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=A3(2);
 CC RA Murphy L., Harris D., Parkhill J., Barrell B.G., Rajandream M.A.;
 CC Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL031184; CAA20173.1; -;
 DR PFAM; PF00464; SHMT; 1.
 DR PROSITE; PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate: One-carbon metabolism.
 FT BINDING 229 229 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 420 AA; 44776 MW; 6BFD07E89AA8CBFC CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 420;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 47 QGSVLT 52
 QY |||||
 14 QGSVLT 19

 CC RESULT 56
 CC ID GLA2_MYCTU STANDARD; PRT; 425 AA.
 CC AC O53615;
 CC DT 15-FEB-2000 (Rel. 39, Created)
 CC DT 15-FEB-2000 (Rel. 39, Last sequence update)
 CC DT 15-FEB-2000 (Rel. 39, Last annotation update)
 CC DE SERINE HYDROXYMETHYLTRANSFERASE 2 (EC 2.1.2.1) (SERINE METHYLASE 2)
 CC (SHMT 2).
 CC GN GLYA2 OR RV0070 OR MTV030.13C.
 CC OS Mycobacterium tuberculosis.
 CC OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=H37RV;
 CC MEDLINE; 98295987.
 CC Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 CC Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 CC Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 CC Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 CC Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 CC Taylor K., Whitehead S., Barrell B.G.;
 CC "Deciphering the biology of Mycobacterium tuberculosis from the
 CC complete genome sequence."
 CC Nature 393:537-544(1998).
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL021428; CAA16251.1; -;
 CC TUBERCULIST; RV0070; -;
 CC PFAM; PF00464; SHMT; 1.
 CC PROSITE; PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate: One-carbon metabolism.
 FT BINDING 230 230 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 425 AA; 45525 MW; 224D195C1D8BF680 CRC64;
 CC
 CC Query Match 5.9%; Score 6; DB 1; Length 425;
 CC Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 47 QGSVLT 52

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 SELAAE 374
|||||

QY 67 SELAAE 72

RESULT 62

ID GLYA_METEX STANDARD; PRT; 434 AA.

AC P50435;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE) (SHMT).

DE (SHMT).

GN GLYA.

OS Methylobacterium extorquens.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Methylobacterium group; Methylobacterium.

[1]

SEQUENCE FROM N.A.

RC STRAIN-AM1 / NCIB 9133;

RX MEDLINE: 95050239.

RA Chistoserdova L.V., Lidstrom M.E.;

RT "Genetics of the serine cycle in Methylobacterium extorquens AM1: cloning, sequence, mutation, and physiological effect of glyA, the gene for serine hydroxymethyltransferase.";

RT J. Bacteriol. 176:6759-6762(1994).

CC -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE. MAY BE REQUIRED FOR BOTH C1 AND C2 METABOLISM.

CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE + H(2)O = TETRAHYDROFOLATE + L-SERINE.

CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS, HORMONES AND OTHER COMPONENTS.

CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

EMBL: L33463; AAA64456.1; -.
PFAM: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW Transferase; Pyridoxal phosphate; One-carbon metabolism.
FT BINDING 242 242 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 434 AA; 46305 MW; EC9599B1E1AE44BC CRC64;

Query Match 5.9%; Score 6; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 QGSVLT 64
|||||

QY 14 QGSVLT 19

RESULT 63

ID SLAP_LACAC STANDARD; PRT; 444 AA.

AC P35829;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE S-LAYER PROTEIN PRECURSOR (SURFACE LAYER PROTEIN) (SA-PROTEIN).

GN SLPA.

OS Lactobacillus acidophilus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;

OC Lactobacillus.

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 343-351 AND 440-444.

RC STRAIN-ATCC 4356;

RX MEDLINE: 94012467.

RA Boot H.J., Kolen C.P.A.M., van Noort J.M., Pouwels P.H.;

RT "S-layer protein of Lactobacillus acidophilus ATCC 4356: purification, expression in Escherichia coli, and nucleotide sequence of the corresponding gene.";

RT J. Bacteriol. 175:6089-6096(1993).

RL

[2]

RN SEQUENCE FROM N.A.

RP STRAIN-ATCC 4356;

RC MEDLINE: 96099308.

RX Boot H.J., Kolen C.P.A.M., Pouwels P.H.;

RT "Identification, cloning, and nucleotide sequence of a silent S-layer protein gene of Lactobacillus acidophilus ATCC 4356 which has extensive similarity with the S-layer protein gene of this species.";

RT J. Bacteriol. 177:7222-7230(1995).

CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.

CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIA IS COVERED BY A S-LAYER WITH HEXAGONAL SYMMETRY.

CC -!- SIMILARITY: SOME, TO THE S-LAYER PROTEIN OF L.BREVIS.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

EMBL: X71412; CAA50535.1; -.
EMBL: X89375; CAA61560.1; -.
DR PIR; A36924; A36924.
KW Signal; Glycoprotein; Cell wall; S-layer.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 444 S-LAYER PROTEIN.
SQ SEQUENCE 444 AA; 46570 MW; 2050732F89099161 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 377 ANIDGT 382
|||||

QY 92 ANIDGT 97

RESULT 64

ID VANS_ENTFA STANDARD; PRT; 447 AA.

AC Q47745;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE SENSOR PROTEIN VANSB (EC 2.7.3.-) (VANCOMYCIN B-TYPE RESISTANCE PROTEIN VANSB) (VANCOMYCIN HISTIDINE PROTEIN KINASE).

GN VANSB.

OS Enterococcus faecalis (Streptococcus faecalis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae; Enterococcus.

OC [1]

RN SEQUENCE FROM N.A.

RP STRAIN-V583;

RX MEDLINE: 96200099.

RA Evers S., Courvalin P.;

RT "Regulation of vanB-type vancomycin resistance gene expression by the Vans(B)-VanR (B) two-component regulatory system in Enterococcus faecalis V583.";

RT J. Bacteriol. 178:1302-1309(1996).

RL

CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM VANS/VANR. ACTIVATES THE TRANSCRIPTION OF VANH, VANA AND VANX IN RESPONSE TO

CC VANCOMYCIN WHICH RESULTS IN VANCOMYCIN RESISTANCE. VANS ACTIVATES
 CC VANK BY PHOSPHORYLATION.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -|- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
 CC KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U35369; AAB05623.1; -.
 DR PFAM; PF00672; DUF5; 1.
 DR PFAM; PF00512; signal; 1.
 KW Sensory transduction; Transferase; Kinase; Transmembrane; Cell wall;
 FT Antibiotic resistance; Phosphorylation.
 TRANSMEM 10 30 POTENTIAL.
 TRANSMEM 137 155 POTENTIAL.
 MOD_RES 233 233 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SEQUENCE 447 AA; 50111 MW; 6471F8A63C7498A2 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 447;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 AQQFVS 35
 |||||
 Qy 77 AQQFVS 82

RESULT 65
 ID AMVB_BACFI STANDARD; PRT; 468 AA.
 AC P96513;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE BETA-AMYLASE PRECURSOR (EC 3.2.1.2) (1,4-ALPHA-D-GLUCAN
 DE MALTOHYDROLASE) (FRAGMENT).
 OS Bacillus firmus.
 CC Bacillus; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen W., He W.B.;
 RT "Nucleotide sequence and characteristics of beta-amylase gene from
 RT Bacillus firmus";
 RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: CATALYZES THE LIBERATION OF MALTOSE FROM 1,4-ALPHA-D
 CC GLUCANS.

CC -|- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN
 CC POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE MALTOSE UNITS FROM THE
 CC NON-REDUCING ENDS OF THE CHAINS.
 CC -|- SIMILARITY: BELONGS TO FAMILY 14 OF GLYCOSYL HYDROLASES
 CC (BETA-AMYLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB000264; BAA19075.1; -.
 DR PFAM; PF01373; Glyco_hydro_14; 1.
 DR PROSITE; PS00506; BETA_AMYLASE_1; 1.
 DR PROSITE; PS00679; BETA_AMYLASE_2; 1.
 KW Hydrolase; Glycosidase; Polysaccharide degradation; Signal.
 FT SIGNAL 1 36 POTENTIAL.
 FT CHAIN 37 >468 BETA-AMYLASE.

FT ACT_SITE 124 124 BY SIMILARITY.
 FT ACT_SITE 198 198 BY SIMILARITY.
 FT NON_TER 468 468
 SQ SEQUENCE 468 AA; 51118 MW; 4B0C11D7FD7316AF CRC64;

Query Match 5.9%; Score 6; DB 1; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 268 TNGGYN 273
 |||||
 Qy 25 TNGGYN 30

RESULT 66
 ID COX1_PLACH STANDARD; PRT; 476 AA.
 AC O92255;

DT 15-FEB-2000 (Rel. 39, Created)
 DT 15-FEB-2000 (Rel. 39, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
 GN COI..

OS Plasmodium chabaudi.

CC Mitochondrion.

CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN [1]

RP SEQUENCE FROM N.A.

RA Tan T.M.C., Noviyanti R., Syafruddin N., Marzuki S., Ting R.C.Y.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

CC -|- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY

CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-

CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE

CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN

CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2

CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3

CC AND COPPER B.

CC -|- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +

CC 4 FERRICYTOCHROME C.

CC -|- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.

CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL

CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.

CC -|- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF014116; AAD01528.1; ALT_SEQ.

DR PFAM; PF00115; COX1; 1.

DR PRINTS; PR01165; CYCOXIDASE1.

DR PROSITE; PS00077; COX1; 1.

DR KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;

DR KW Respiratory chain; Inner membrane.

FT METAL 66 66 IRON (HEME A) (PROBABLE).

FT METAL 246 246 COPPER B (PROBABLE).

FT METAL 250 250 COPPER B (PROBABLE).

FT METAL 295 295 COPPER B (PROBABLE).

FT METAL 296 296 COPPER B (PROBABLE).

FT METAL 382 382 IRON (HEME A3) (PROBABLE).

FT METAL 384 384 IRON (HEME A) (PROBABLE).

FT METAL 476 AA; 52783 MW; 83DA0694884F33BA CRC64;

SEQ SEQUENCE

Query Match 5.9%; Score 6; DB 1; Length 476;
 Best Local Similarity 100.0%; Pred. No. 8.75e+00;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 93 GSSELA 98

|||||

Qy 65 GSSELA 70

```

RESULT 67
ID COX1_PLABE STANDARD; PRT; 476 AA.
AC O99252;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
GN COI.
OS Plasmodium berghei.
OG Mitochondrion.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA Tan T.M.C., Noviyanti R., Syafrudi N., Marzuki S., Ting R.C.Y.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
AND COPPER B.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
4 FERRICYTOCHROME C.
CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF014115; AAD01525.1; ALT_SEQ.
DR PFAM; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
Respiratory chain; Inner membrane.
FT METAL 66 66 IRON (HEME A) (PROBABLE).
FT METAL 246 246 COPPER B (PROBABLE).
FT METAL 250 250 COPPER B (PROBABLE).
FT METAL 295 295 COPPER B (PROBABLE).
FT METAL 296 296 COPPER B (PROBABLE).
FT METAL 382 382 IRON (HEME A3) (PROBABLE).
FT METAL 384 384 IRON (HEME A) (PROBABLE).
SQ SEQUENCE 476 AA; 52728 MW; 69016E0EE61091C0 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 93 GSSELA 98
QY [11111]
65 GSSELA 70

RESULT 68
ID MKK1_YEAST STANDARD; PRT; 508 AA.
AC P32490;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE PROTEIN KINASE MKK1/SSP32 (EC 2.7.1.-).
GN MKK1 OR SSP32 OR YOR231W OR OS095.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

```

```

Saccharomycetaceae; Saccharomyces.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 93233668.
RA Irie K., Takase M., Lee K.S., Levin D.E., Araki H., Matsumoto K.,
RA Oshima Y.;
RT "MKK1 and MKK2, which encode Saccharomyces cerevisiae
RT mitogen-activated protein kinase-kinase homologs, function in the
RT pathway mediated by protein kinase C.";
RL Mol. Cell. Biol. 13:3076-3083(1993).
[2]
RN RP SEQUENCE FROM N.A.
RX STRAIN=S288C / FY1679;
RX MEDLINE; 97127829.
RA Boyer J., Michaux G., Fairhead C., Gaillon L., Dujon B.;
RT "Sequence and analysis of a 26.9 kb fragment from chromosome XV of
RT the yeast Saccharomyces cerevisiae.";
RL Yeast 12:1575-1586(1996).
CC -!- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
INVOLVE THE KINASE PKC1 THAT MAY ACT ON THE BCK1 KINASE THAT THEN
PHOSPHORYLATES MKK1 AND MKK2 WHICH THEMSELVES PHOSPHORYLATE THE
MKK1 KINASE.
CC -!- FUNCTION: KINASE INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT
INVOLVE THE PROTEIN KINASES ENCODED BY PKC1, BCK1, AND MPK1.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
MAP KINASE KINASE SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D13001; BAA02364.1; -.
DR EMBL; Z75139; CAA99451.1; -.
DR PIR; S30772; S30772.
DR PIR; A48069; A48069.
DR HSSP; P11362; IFGI.
DR SGD; L0001117; MKK1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
ATP-binding; Phosphorylation.
FT DOMAIN 221 488 PROTEIN KINASE.
FT NP_BIND 227 235 ATP (BY SIMILARITY).
FT BINDING 250 250 ATP (BY SIMILARITY).
FT ACT_SITE 349 349 BY SIMILARITY.
FT MOD_RES 377 377 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 381 381 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 508 AA; 56720 MW; 4FE62CDD185CDA2 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 196 IDLSNV 201
QY [11111]
34 IDLSNV 39

RESULT 69
ID CD39_HUMAN STANDARD; PRT; 510 AA.
AC P49961;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VASCULAR ATP-DIPHOSPHOHYDROLASE (EC 3.6.1.5) (ATPDASE) (LYMPHOID CELL

```


DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 61.5 KDA PROTEIN U1937B.
GN U1937B OR B1937_F1_4.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Mycobacteriaceae; Mycobacterium.
CC Actinomycetales; Corynebacterineae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO SYNECHOCYSTIS PCC 6803 SLL0335, SOME TO
CC M. TUBERCULOSIS RV2567.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00016; AAA17160.1; -;
CC Hypothetical protein.
FT DOMAIN 538 554 GLN-RICH
SQ SEQUENCE 558 AA; 61522 MW; 4472E500100ABFA6 CRC64;

Query Match 5.9%; Score 6; DB 1; Length 558;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 29 GGYNTS 34
| | | | |
QY 27 GGYNTS 32

RESULT 74
ID HEMA_TAPIL STANDARD; PRT; 566 AA.
AC P13102;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HAI CHAIN;
DE HEMAGGLUTININ HAZ CHAIN].
GN HA.
OS Influenza A virus (strain A/Pilot whale/Maine/328/84).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza virus A and B group.
RN [1]
RP SEQUENCE FROM N.A.
RA Chambers T.M., Yamnikova S., Kawaoka Y., Lvov D.K., Webster R.G.;
RT "Antigenic and molecular characterization of subtype H13
hemagglutinin of influenza virus.";
RL Virology 172:180-188(1989).
CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMER IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M26091; AAA43215.1; -;
CC PIR; B32664; HMIIVT2.
CC HSSP; P03437; 2HMG.
CC PFAM; PF00509; Hemagglutinin; 1.
CC PRINTS; PR00329; HEMAGGLUTIN12.
CC PRINTS; PR00330; HEMAGGLUTIN1.
CC PRINTS; PR00331; HEMAGGLUTIN2.

DR PRINTS; PR00331; HEMAGGLUTIN2.
KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 342 HEMAGGLUTININ HAI CHAIN.
FT CHAIN 344 566 HEMAGGLUTININ HAZ CHAIN.
FT CARBOHYD 29 29 POTENTIAL.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 182 182 POTENTIAL.
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 305 305 POTENTIAL.
FT CARBOHYD 488 488 POTENTIAL.
FT CARBOHYD 497 497 POTENTIAL.
SQ SEQUENCE 566 AA; 63049 MW; 74FF5A9860B8E59E CRC64;

Query Match 5.9%; Score 6; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 TSSIDL 50
| | | | |
QY 31 TSSIDL 36

RESULT 75
ID HEMA_IAGU2 STANDARD; PRT; 566 AA.
AC P13103;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HEMAGGLUTININ PRECURSOR [CONTAINS: HEMAGGLUTININ HAI CHAIN;
DE HEMAGGLUTININ HAZ CHAIN].
GN HA.
OS Influenza A virus (strain A/Gull/Maryland/704/77).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza virus A and B group.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91220697.
RA Nobusawa E., Aoyama T., Kato H., Suzuki Y., Tateno Y., Nakajima K.;
RT "Comparison of complete amino acid sequences and receptor-binding
properties among 13 serotypes of hemagglutinins of influenza A
viruses.";
RL Virology 182:475-485(1991).
CC [2]
RN SEQUENCE FROM N.A.
RP MEDLINE; 89370299.
RA Chambers T.M., Yamnikova S., Kawaoka Y., Lvov D.K., Webster R.G.;
RT "Antigenic and molecular characterization of subtype H13
hemagglutinin of influenza virus.";
RL Virology 172:180-188(1989).
CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMER IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90308; BAA14338.1; -;
CC EMBL; M26090; AAA43214.1; -;
CC PIR; C39987; HMIIVGM.
CC PIR; A32664; HMIIVT1.
CC PIR; G33157; G33157.
CC HSSP; P03437; 2HMG.
CC PFAM; PF00509; Hemagglutinin; 1.
CC PRINTS; PR00329; HEMAGGLUTIN12.
CC PRINTS; PR00330; HEMAGGLUTIN1.
CC PRINTS; PR00331; HEMAGGLUTIN2.

KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 342 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 344 566 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 29 29 POTENTIAL.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 182 182 POTENTIAL.
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 305 305 POTENTIAL.
FT CARBOHYD 488 488 POTENTIAL.
FT CARBOHYD 497 497 POTENTIAL.
FT CONFLICT 80 80 V -> G (IN REF. 2).
FT CONFLICT 172 172 T -> N (IN REF. 2).
SQ SEQUENCE 566 AA; 63294 MW; 0965B35148F2215F CRC64;

Query Match 5.9%; Score 6; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 8.75e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

45 TSSIDL 50
|||||
QY 31 TSSIDL 36

Search completed: Wed Aug 16 09:55:49 2000
Job time : 45 secs.

